

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date  
15 September 2005 (15.09.2005)

PCT

(10) International Publication Number  
**WO 2005/084430 A1**

(51) International Patent Classification<sup>7</sup>: **A01K 67/027**

(US). LEAVITT, Markley, C. [US/US]; 1140 Latham Drive, Watkinsville, GA 30677 (US).

(21) International Application Number:

PCT/US2005/006072

(74) Agent: YESLAND, Kyle; Legal Department, AviGenics, Inc., 111 Riverbend Road, Athens, GA 30605 (US).

(22) International Filing Date: 28 February 2005 (28.02.2005)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

10/790,455 1 March 2004 (01.03.2004) US  
10/811,136 26 March 2004 (26.03.2004) US  
10/940,315 14 September 2004 (14.09.2004) US

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

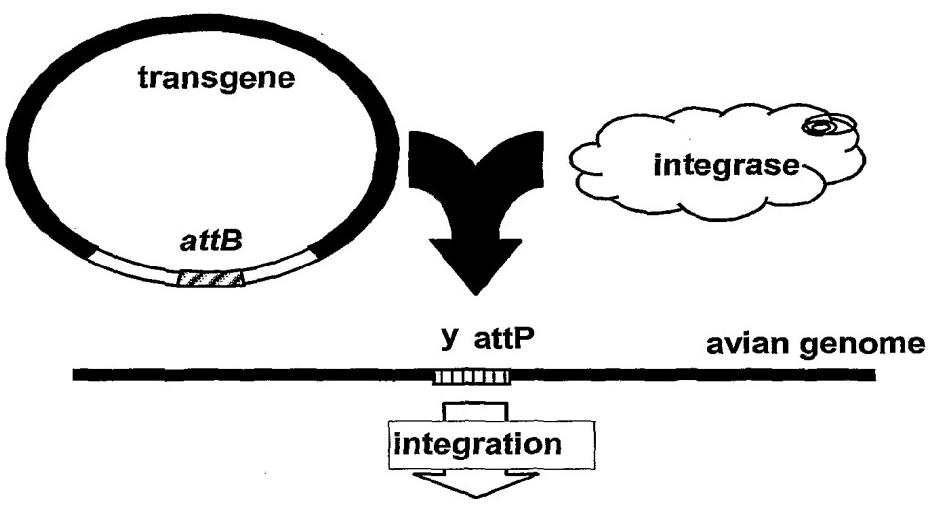
(71) Applicant (for all designated States except US): AVIGENICS, INC [US/US]; Legal Department, 111 Riverbend Road, Athens, GA 30605 (US).

Published:

— with international search report

[Continued on next page]

(54) Title: GENERATION OF TRANSGENIC AVIANS



WO 2005/084430 A1

(57) Abstract: The invention includes transchromosomal avians and transchromosomal avian cells and methods for the introduction of artificial chromosomes into the genome of avians and avian cells.



— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

## GENERATION OF TRANSGENIC AVIANS

The present application claims priority to U.S. Patent Application  
5 10/940,315, filed September 14, 2004; U.S. Patent Application No. 10/811,136,  
filed March 26, 2004 and U.S. Patent Application No. 10/790,455, filed March 1,  
2004.

**Field of the Invention**

10 The present invention relates to the field of biotechnology, and more specifically to the field of genome modification. Disclosed herein are compositions including chromosomes and vectors, and methods of use thereof, for the generation of genetically transformed cells and animals.

15 **Background**

Transgenic technology to convert animals into "bioreactors" for the production of specific proteins or other substances of pharmaceutical interest (Gordon et al, 1987, Biotechnology 5: 1183-1187; Wilmut et al, 1990, Theriogenology 33: 113-123) offers significant advantages over more conventional methods of protein production by gene expression. For example, recombinant nucleic acid molecules have been engineered and incorporated into transgenic animals so that an expressed heterologous protein may be joined to a protein or peptide that allows secretion of the transgenic expression product into milk or urine, from which the protein may then be recovered.

25 Another system useful for heterologous protein production is the avian reproductive system. The production of an avian egg begins with formation of a large yolk in the ovary of the hen. The unfertilized oocyte or ovum is positioned on top of the yolk sac. After ovulation the ovum passes into the infundibulum of the oviduct where it is fertilized, if sperm are present, and then moves into the magnum of the oviduct which is lined with tubular gland cells. These cells secrete the egg-white proteins, including ovalbumin, lysozyme, ovomucoid, conalbumin and ovomucin into the lumen of the magnum where they are deposited onto the avian embryo and yolk. The hen oviduct offers outstanding potential as a protein

bioreactor because of the high levels of protein production, the promise of proper folding and post-translation modification of the target protein, the ease of product recovery, and the relatively short developmental period of chickens.

One method for creating permanent genomic modification of a eukaryotic  
5 cell is to integrate an introduced DNA into an existing chromosome. Retroviruses have so far proven to be the method of choice for efficient integration. However, retroviral integration is directed to a number of insertion sites within the recipient genome so that positional variation in heterologous gene expression can be evident. Unpredictability as to which insertion site is targeted introduces an undesirable  
10 lack of control over the procedure. An additional limitation of the use of retroviruses is that the size of the nucleic acid molecule encoding the virus and heterologous sequences may be limited to about 8 kb. In addition, retroviruses may include undesirable features such as splice sites. Although wild-type adeno-associated virus (AAV) often integrates at a specific region in the human genome,  
15 replication deficient vectors derived from AAV do not integrate site-specifically possibly due to the deletion of the toxic rep gene. In addition, homologous recombination produces site-specific integration, but the frequency of such integration usually is typically low.

An alternative method for delivering a heterologous nucleic acid into the  
20 genome is the use of a site-specific enzymes that can catalyze the insertion of nucleic acids into chromosomes. These enzymes recognize relatively short unique nucleic acid sequences that serve for both recognition and recombination. Examples include Cre (Sternberg & Hamilton, 1981, J. Mol. Biol. 150: 467-486, 1981), Flp (Broach et al, 1982, Cell 29: 227-234, 1982) and R (Matsuzaki et al,  
25 1990, J. Bact. 172: 610-618, 1990).

A novel class of phage integrases that includes the integrase from the phage phiC31 can mediate highly efficient integration of transgenes in mammalian cells both in vitro and in vivo (Thyagarajan et al, Mol. Cell Biol. 21: 3926-3934, 2001). Constructs and methods of using recombinase to integrate heterologous DNA into  
30 a plant, insect or mammalian genome are described by Calos in U.S. Patent Serial No. 6,632,672, the disclosure of which is incorporated in its entirety herein by reference.

The phiC31 integrase is a member of a subclass of integrases, termed serine recombinases, that include, for example, R4 and TP901-1. Unlike the phage lambda integrases, which belong to a tyrosine class of recombinases, the serine integrases do not require cofactors such as integration host factor. The phiC31 5 integrase normally mediates integration of the phiC31 bacteriophage into the genome of Streptomyces via recombination between the attP recognition sequence of the phage genome and the attB recognition sequence within the bacterial genome. When a plasmid is equipped with a single attB site, phiC31 integrase will detect and mediate crossover between the attB site and a pseudo-attP site within the 10 mammalian genome. Such pseudo-attP integration sites have now been identified in the mouse and human genomes. If the heterologous DNA is in a circular or supercoiled form, the entire plasmid becomes integrated with attL and attR arms flanking the nucleic acid insert. PhiC31 integrase is not able to mediate the 15 integration into genomic DNA of sequences bearing attP sites.

Integration mediated by certain integrases, such as PhiC31 integrase-mediated integration, results in the destruction of the recognition or recombination sites themselves so that the integration reaction is irreversible. This will bypass the primary concern inherent with other recombinases, i.e., the reversibility of the 20 integration reaction and excision of the inserted DNA.

Another method for the stable introduction of heterologous nucleic acid (e.g., large heterologous nucleic acids) into a genome is by the use of an artificial chromosome. Artificial chromosomes for expression of heterologous genes in yeast are available, but artificial chromosomes being delivered to avians has not 25 previously been achieved.

Therefore, it is an object of the invention to produce transgenic animals with large nucleic acid segments integrated into their genome and to provide avians which include an artificial chromosome in their genome.

#### Summary of the Invention

Integration of a transgene into a defined chromosomal site is useful to improve the predictability of expression of the transgene, which is particularly advantageous when creating transgenic vertebrate animals such as, transgenic 30

avians. Transgenesis by methods that randomly insert a transgene into a genome are often inefficient since the transgene may not be expressed at the desired levels or in desired tissues.

The present invention relates to methods of modifying the genome of vertebrate cells (e.g., production of transgenic vertebrates) and to such cells with modified genomes and their progeny. In one embodiment, the methods provide for introducing into vertebrate cells a first recombination site such that the recombination site is inserted into the vertebrate cell genome. Typically, in such embodiments, the genome does not normally include this first recombination site prior to the recombination site introduction. Methods of the invention may also include introducing a nucleotide sequence comprising a second recombination site and a sequence of interest such as a coding sequence into the vertebrate cell or progeny of the vertebrate cell. The nucleotide sequence comprising the second recombination site and the sequence of interest such as a coding sequence may be introduced into the vertebrate cell before, at about the same time as or after the introduction of the first recombination site. Additionally, the present methods may include introducing into the vertebrate cell or progeny cell thereof a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and the sequence of interest proximal to the first recombination site. For example, the nucleotide sequence comprising the second recombination site and the sequence of interest may be inserted adjacent to or internally in the first recombination site. In one very useful embodiment, the first recombination site and/or the nucleotide sequence comprising the second recombination site and the sequence of interest are stably incorporated into the genome of the cell.

The present invention contemplates the genomic modification of any useful vertebrate cells including, but not limited to, avian cells. Examples of cells which may have their genomes modified in accordance with the present invention include, without limitation, reproductive cells including sperm, ova and embryo cells and nonreproductive cells such as tubular gland cells.

The present invention also relates to methods of producing transgenic vertebrate animals and to the transgenic animals produced by the methods and to their transgenic progeny or descendants. The invention also includes the transgenic

cells included in or produced by the transgenic vertebrate animals. Examples of such cells include, without limitation, germ line cells, ova, sperm cells and protein producing cells such as tubular gland cells. In one useful embodiment, the transgenic vertebrate animals of the invention are transgenic avians. Transgenic  
5 avians of the invention may include, without limitation, chickens, turkeys, ducks, geese, quail, pheasants, parrots, finches, hawks, crows or ratites including ostrich, emu or cassowary.

In accordance with the present invention, methods of producing transgenic vertebrate animals can include introducing into an embryo of a vertebrate animal a  
10 first recombination site such that the recombination site is present in sperm or ova of a mature vertebrate animal developed from the embryo. In one useful embodiment, the embryo does not normally include the first recombination site in its genome prior to the recombination site introduction. The methods may also include introducing a nucleotide sequence comprising a second recombination site  
15 and a sequence of interest such as a coding sequence into the embryo of the vertebrate animal. The first recombination site and/or the nucleotide sequence comprising the second recombination site and a sequence of interest may be introduced into the embryo of the vertebrate animal before the embryo is fertilized (i.e., when an ovum), at about the same time as introduction of the sperm into the  
20 ovum or after fertilization.

The methods can also include introducing the nucleotide sequence comprising a second recombination site and a sequence of interest into an ovum or a sperm of a mature vertebrate animal developed from the embryo (or its descendants) into which the first recombination site was introduced. In one  
25 embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum from the mature vertebrate animal before the ovum is fertilized. In another embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum at about the time of fertilization. In one particularly  
30 useful embodiment, the nucleotide sequence comprising a second recombination site and a sequence of interest is introduced into the ovum after the ovum is fertilized (when an embryo).

The methods may include, upon addition of the nucleotide sequence comprising a second recombination site and a sequence of interest to an embryo, ovum or sperm, introducing into the embryo, ovum or sperm, a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and the sequence of interest proximal to the first recombination site. For example, the nucleotide sequence comprising the second recombination site and the sequence of interest may be inserted adjacent to or internally in the first recombination site. In one useful embodiment, the methods include introducing into an embryo comprising the first recombination site in its genome, a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and the sequence of interest proximal to the first recombination site.

In one useful embodiment, these methods include fertilizing an ovum with sperm comprising the first recombination site. The methods can include also introducing into the ovum a nucleotide sequence comprising a second recombination site and a sequence of interest such as a coding sequence and a substance which facilitates insertion of the nucleotide sequence comprising the second recombination site and sequence of interest proximal to (e.g., adjacent to or internally in) the first recombination site. It is contemplated that the nucleotide sequence comprising a second recombination site and a sequence of interest may be introduced into the ovum before or after fertilization by the sperm or at about the same time as fertilization.

In one very useful embodiment of the methods disclosed herein, the nucleotide sequence comprising the second recombination site and the sequence of interest is stably incorporated into the genome of the embryo, ovum or sperm.

The methods disclosed herein typically eventually include exposing a fertilized ovum to conditions which lead to the development of a viable transgenic vertebrate animal.

In one embodiment, the nucleotide sequence of interest includes an expression cassette. Optionally, the nucleotide sequence of interest may include a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP-encoding gene, and the like.

Typically, in accordance with methods known in the art or methods disclosed herein, the embryo of the vertebrate animal or fertilized ovum of a mature vertebrate animal of the invention is exposed to conditions which lead to the development of a viable transgenic vertebrate animal.

5        Embryos that are useful in the present methods include, without limitation, stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryos.

In one embodiment, the nucleotide sequence included with the second recombination site of interest is a coding sequence. The nucleotide sequence of 10 interest included with the second recombination site can be of any useful size. For example, and without limitation, the nucleotide sequence of interest may be from about 0.1 kb to about 10 mb, for example, about 1 kb to about 1 mb. In one embodiment, the nucleotide sequence of interest is about 5 kb to about 5 mb in size, for example, about 5 kb to about 2 mb, e.g., about 8 kb to about 1 mb. In one 15 embodiment, the nucleotide sequence of interest is about 0.5 kb to about 500 kb.

The first recombination site and/or the nucleotide sequence which includes the second recombination site and a sequence of interest such as a coding sequence 20 may be introduced into cells, embryos (i.e., fertilized ova) or sperm by any useful method. These useful methods include, without limitation, cell fusion, lipofection, transfection, microinjection, calcium phosphate co-precipitation, electroporation, protoplast fusion, particle bombardment and the like. In addition, the first 25 recombination site or nucleotide sequence comprising the second recombination site and the sequence of interest may be introduced into cells, embryos, ova or sperm in the presence of a cationic polymer such as PEI and/or other substances disclosed elsewhere herein or known in the art.

In one embodiment, recombination sites employed in the present invention are isolated from bacteriophage and/or bacteria. For example, the recombination sites may be attP sites or attB sites.

The substance which facilitates insertion of the second recombination site 30 and a sequence of interest may be an enzyme. In one embodiment, the substance is a site specific recombinase. In one useful embodiment, the substance which facilitates insertion of the nucleotide sequence is nucleic acid, for example, DNA

or RNA. The DNA or RNA may include modified nucleosides as described elsewhere herein or are known to those of skill in the art. In one embodiment, modified nucleosides are employed to extend the half-life of RNA or DNA molecules employed in the present invention. For example, it may be desirable to 5 extend the half life of the RNA or DNA molecules in the presence of a cellular environment. In one useful embodiment, the nucleic acid encodes an enzyme such as a site specific recombinase.

Nonlimiting examples of site specific recombinases which may be employed herein either as protein or encoded by nucleic acid include serine 10 recombinases and tyrosine recombinases. Examples of serine recombinases which may be employed include, without limitation, EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1,  $\Phi$ 370.1,  $\Phi$ 105,  $\Phi$ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4,  $\Phi$ Rv1, Y4bA and Bja serine 15 recombinases.

In one embodiment of the invention, the present methods include introducing an integration host factor into a cell (e.g., an embryo) to facilitate genomic integration. Such integration host factors may be particularly useful when employing certain substances such as tyrosine recombinases as disclosed herein.

20 The nucleotide sequence of interest may include a coding sequence. The coding sequence may encode any useful protein. In one useful embodiment, the sequence of interest encodes a pharmaceutical or therapeutic substance. The invention contemplates the production of any useful protein based pharmaceutical or therapeutic substance. Examples of pharmaceutical or therapeutic substances 25 include without limitation at least one of a light chain or a heavy chain of an antibody (e.g., a human antibody) or a cytokine. In one embodiment, the pharmaceutical or therapeutic composition is interferon, erythropoietin, or granulocyte-colony stimulating factor. In one embodiment, the transgenic animal is an avian and the sequence of interest encodes a polypeptide present in eggs 30 produced by the avian.

In one embodiment, integrases such as phage integrases, for example, serine recombinases, such as the integrase from phage phiC31, can mediate the efficient

integration of transgenes into target cells both in vitro and in vivo. In one embodiment, when a plasmid is equipped with a single attB site, the integrase detects attP homologous sequences, termed pseudo-attP sites, in a target genome and mediates crossover between the attB site and a pseudo attP site.

5       In one embodiment, once delivered to a recipient cell, for example, an avian cell, the phiC31 integrase mediates recombination between the att site within the nucleic acid molecule and a bacteriophage attachment site within the genomic DNA of the cell. Both att sites are disrupted and the nucleic acid molecule, with partial att sequences at each end, is stably integrated into the genome attP site. The  
10      phiC31 integrase, by disrupting the att sites of the incoming nucleic acid and of the recipient site within the cell genome can preclude any subsequent reverse recombination event that would excise the integrated nucleic acid and reduce the overall efficiency of stable incorporation of the heterologous nucleic acid.

Following delivery of the nucleic acid molecule and a source of integrase activity into a cell population and integrase-mediated recombination, the cells may be returned to an embryo. In the case of avians, late stage blastodermal cells may be returned to a hard shell egg, which is resealed for incubation until hatching. Stage I embryos may be directly microinjected with the polynucleotide and source of integrase activity, isolated, transfected and returned to a stage I embryo which is reimplemented into a hen for further development. Additionally, the transfected cells may be maintained in culture in vitro.  
15  
20

The present invention provides novel methods and recombinant polynucleotide molecules for transfecting and integrating a heterologous nucleic acid molecule into the genome of a cell of a vertebrate animal, such as an avian.  
25      Certain methods of the invention provide for the delivery to a cell population a first nucleic acid molecule that comprises a region encoding a recombination site, such as a bacterial recombination site or a bacteriophage recombination site. In one embodiment, a source of integrase activity is also delivered to the cell and can be in the form of an integrase-encoding nucleic acid sequence and its associated promoter or as a region of a second nucleic acid molecule that may be co-delivered with the polynucleotide molecule. Alternatively, integrase protein itself can be delivered directly to the target cell.  
30

The recombinant nucleic acid molecules of the present invention may further comprise a heterologous nucleotide sequence operably linked to a promoter so that the heterologous nucleotide sequence, when integrated into the genomic DNA of a recipient cell, can be expressed to yield a desired polypeptide. The 5 nucleic acid molecule may also include a second transcription initiation site, such as an internal ribosome entry site (IRES), operably linked to a second heterologous polypeptide-encoding region desired to be expressed with the first polypeptide in the same cell.

The present invention provides modified isolated artificial chromosomes 10 useful as vectors to shuttle transgenes or gene clusters into a genome of an avian. By delivery of the modified chromosome to a recipient cell, the target cell, and progeny thereof, become trisomic or transchromosomal. The additional chromosome will typically not affect the subsequent development of the recipient cell and/or embryo, nor interfere with the reproductive capacity of an adult bird 15 developed from such cells or embryos. The chromosome will also be stable within the genome of the cells of the adult bird or within isolated avian cells. The invention provides methods to isolate a population of chromosomes for delivery into embryos or early cells of avians, for example, chickens.

The methods can include inserting a lac-operator sequence into an isolated 20 chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. The lac operator region is typically a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. A recombinant DNA molecule is constructed that includes an identified region of the target chromosome, a recombination site such as attB or attP, and the lac-operator 25 concatamer. The recombinant molecule is delivered to an avian cell, and homologous recombination will integrate the heterologous polynucleotide and the lac-operator concatamer into the targeted chromosome. A tag-polypeptide, such as the GPF-lac-repressor fusion protein, binds to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged 30 mitotic chromosome can be isolated using, for instance, flow cytometry.

Among other things, the present invention relates to transchromosomal avians. In a particular aspect, the invention provides for G0 transchromosomal

avians (e.g., germline chimeric transchromosomal avians) which can produce germline transchromosomal offspring (e.g., G1 and G2 germline transchromosomal offspring).

Examples of avians which are contemplated for use herein include, without limitation, chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

In one useful aspect, the artificial chromosome employed herein includes a centromere. Any useful centromere may be employed in the present invention including, without limitation, centromeres from insects, mammals or avians.

10 In one particularly useful embodiment, the artificial chromosomes used herein include a heterologous nucleotide sequence. The nucleotide sequence may be heterologous to the avian and/or heterologous to the artificial chromosome. In one useful embodiment, the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance. In addition, the heterologous nucleotide sequence may include a gene expression controlling region. Any useful gene expression controlling region may be employed in the invention. For example, and without limitation, the gene expression controlling region may include a lysozyme promoter, an ovomucin promoter, a conalbumin promoter, an ovomucoid promoter and/or an ovalbumin promoter or functional portions thereof. See, for example,  
15 US Patent Application No. 10/114,739, filed April 1, 2002; US Patent Application No. 10/856,218, filed May 28, 2004 and US Patent Application No. 10/733,042, filed December 11, 2003. The disclosure of each of these patent applications is incorporated herein by reference in its entirety. In one useful embodiment, the product of the heterologous nucleotide sequence (e.g., therapeutic substance) is  
20 delivered to the avian egg (e.g., the egg white) during production of the egg in the avian. The invention also includes the eggs produced by the avians produced by  
25 these methods and other methods disclosed herein.

One useful aspect of the invention relates to methods of producing transchromosomal avians. In one embodiment, the methods include substantially purifying a chromosome followed by introducing the purified chromosome into an avian embryo and thereafter maintaining the embryo under conditions suitable for the embryo to develop and hatch as a chick. In one embodiment, the methods

include inserting a heterologous nucleotide sequence into the chromosome before substantially purifying the chromosome. In one embodiment, the chromosome is introduced into the avian embryo by microinjection; however, any useful method to introduce the chromosome into the avian embryo is within the scope of the present invention.

It is contemplated that the chromosome may be introduced into the embryo by delivering the chromosome to an avian cell before or after fertilization. For example, the chromosome may be introduced into an ovum or a sperm before fertilization. In another example, the chromosome is introduced into a cell of an embryo (e.g., stage I to stage XII embryo). In one embodiment, the chromosome is introduced into an early stage embryo, for example, and without limitation, a stage I embryo. In one embodiment, the chromosome is introduced into a germinal disc.

The methods provide for the introduction of any useful number of chromosomes into the avian embryo in order to produce a transchromosomal avian. For example, and without limitation, between 1 and about 10,000 chromosomes may be introduced into the embryo. In another example, between 1 and about 1,000 chromosomes may be introduced into the embryo.

The invention also provides for transchromosomal avian cells wherein the artificial chromosome includes a nucleotide sequence which encodes a therapeutic substance. The cells may be isolated from transchromosomal avians and thereafter grown in culture. The invention also contemplates the production of the transchromosomal avian cells by stable introduction of the artificial chromosome into cultured avian cells. Any useful method may be employed for the introduction of the artificial chromosome into the cultured cells including, without limitation, lipofection or microinjection.

Another aspect of the present invention is a cell, for example, an avian cell, genetically modified with a transgene vector by the methods of the invention. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. One useful cell line

for the delivery and integration of a transgene comprises a heterologous attP site that can increase the efficiency of integration of a polynucleotide by an integrase, such as phiC31 integrase and, optionally, a region for expressing the integrase.

Another aspect of the present invention is methods of expressing a  
5 heterologous polypeptide in a cell by stably transfected a cell by using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of a transcriptional regulatory region.

Yet another aspect of the present invention concerns transgenic vertebrate  
10 animals, such as birds, for example chickens, comprising a recombinant nucleic acid molecule and which may (though optionally) express a heterologous gene in one or more cells in the animal. For example, in the case of avians, embodiments of the methods for the production of a heterologous polypeptide by the avian tissue involve providing a suitable vector and introducing the vector into embryonic  
15 blastodermal cells containing an attP site together with an integrase, for example, a serine recombinase such as phiC31 integrase, so that the vector can integrate into the avian genome at the attP site which has been engineered into the cell genome. A subsequent step may involve deriving a mature transgenic avian from the transgenic blastodermal cells by transferring the transgenic blastodermal cells to an  
20 embryo, such as a stage X embryo (e.g., an irradiated stage X embryo), and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. In one embodiment, sperm from a G0 bird positive for the transgene is used to inseminate a chicken giving rise to a fully transgenic G1 generation.

25 One approach may be to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult animal. The resulting animal is then grown to maturity.

In the transgenic vertebrate of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental  
30 stages utilizing, for example, trans-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern.

Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences. By 5 inserting an integration site such as attP into the genome, it is believed that expression of an integrated coding sequence will be much more predictable.

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth 10 hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β-casein may be produced by the present methods. In one embodiment, proteins are expressed in the oviduct and deposited in eggs of avians, such as chickens, according to the invention. The present invention includes these eggs and these proteins.

15 The present invention also includes methods of producing transgenic vertebrate animals, for example, transgenic chickens, which employ the use of integrase, cationic polymers and/ nuclear localization signals. The present invention also includes the transgenic vertebrate animals, such as the avians, produced by these methods and other methods disclosed herein. The invention also 20 includes the eggs produced by the transgenic avians produced by these methods and other methods disclosed herein.

In one embodiment, the methods of the invention include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; and 3) a cationic polymer. Such methods provide for an increased efficiency of transgenic 25 avian production relative to identical methods without the cationic polymer.

In another embodiment, the methods include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; and 3) a nuclear localization signal. Such methods provide for an increased efficiency of transgenic animal, for example, avian, production relative to identical methods without the 30 nuclear localization signal.

In another embodiment, the methods include introducing into a cell: 1) a nucleic acid comprising a transgene; 2) an integrase activity; 3) a cationic polymer;

and 4) a nuclear localization signal. Such methods provide for an increased efficiency of transgenic vertebrate animal production relative to identical methods without the cationic polymer or without the nuclear localization signal.

In one embodiment, the cell is a cell of an embryo, for example, an avian embryo. In one embodiment, the cell is a cell of an early stage avian embryo comprising a germinal disc. The avian cell may be, for example, a cell of a stage I avian embryo, a cell of a stage II avian embryo, a cell of a stage III avian embryo, a cell of a stage IV avian embryo, a cell of a stage V avian embryo, a cell of a stage VI avian embryo, a cell of a stage VII avian embryo, a cell of a stage VIII avian embryo, a cell of a stage IX avian embryo, a cell of a stage X avian embryo, a cell of a stage XI avian embryo or a cell of a stage XII avian embryo. In one particularly useful embodiment, the avian cell is a cell of a stage X avian embryo. In another useful embodiment, the avian cell is a cell of a stage I avian embryo.

The methods provide for the introduction of nucleic acid into the avian cell by any suitable technique known to those of skill in the art. For example, the nucleic acid may be introduced into the avian cell by microinjecting, transfection, electroporation or lipofection. In one particularly useful embodiment, the introduction of the nucleic acid is accomplished by microinjecting.

The nucleic acid which includes a transgene may be DNA or RNA or a combination of RNA and DNA. The nucleic acid may comprise a single strand or may comprise a double strand. The nucleic acid may be a linear nucleic acid or may be an open or closed circular nucleic acid and may be naturally occurring or synthetic.

Integrase activity may be introduced into the cell, such as an avian cell, in any suitable form. In one embodiment, an integrase protein is introduced into the cell. In another embodiment, a nucleic acid encoding an integrase is introduced into the cell. The nucleic acid encoding the integrase may be double stranded DNA, single stranded DNA, double stranded RNA, single stranded RNA or a single or double stranded nucleic acid which includes both RNA and DNA. In one particularly useful embodiment, the nucleic acid is mRNA. Integrase activity may be introduced into the cell by any suitable technique. Suitable techniques include those described herein for introducing the nucleic acid encoding a transgene into a

cell. In one useful embodiment, the integrase activity is introduced into the cell with the nucleic acid encoding the transgene. For example, the integrase activity may be introduced into the cell in a mixture with the nucleic acid encoding the transgene.

5       In one embodiment, a nuclear localization signal (NLS) is associated with the nucleic acid which includes a transgene. For example, the NLS may be associated with the nucleic acid by a chemical bond. Examples of chemical bonds by which an NLS may be associated with the nucleic acid include an ionic bond, a covalent bond, hydrogen bond and Van der Waal's force. In one particularly  
10 useful embodiment, the nucleic acid which includes a transgene is associated with an NLS by an ionic bond. NLS may be introduced into the cell by any suitable technique. Suitable techniques included those described herein for introducing the nucleic acid encoding a transgene into a cell. In one useful embodiment, the NLS is introduced into the cell with the nucleic acid encoding the transgene. For  
15 example, the NLS may be introduced into the cell while associated with the nucleic acid encoding the transgene.

Cationic polymers may be employed to facilitate the production of transgenic vertebrate animals such as avians. For example, the cationic polymers may be employed in combination with integrase and/or NLS. Any suitable cationic  
20 polymer may be used. For example, and without limitation, one or more of polyethylenimine, polylysine, DEAE-dextran, starburst dendrimers and starburst polyamidoamine dendrimers may be used. In a particularly useful embodiment, the cationic polymer includes polyethylenimine. The cationic polymer may be introduced into the cell by any suitable technique. Suitable techniques included  
25 those described herein for introducing the nucleic acid encoding a transgene into a cell. In one useful embodiment, the cationic polymer is introduced into the cell in a mixture with the nucleic acid encoding the transgene. For example, the cationic polymer may be introduced into the avian cell while associated with the nucleic acid encoding the transgene.

30       In one particularly useful embodiment of the invention, the transgene includes a coding sequence which is expressed in a cell of the transgenic vertebrate animal, for example, a transgenic avian, producing a peptide or a polypeptide (e.g.,

a protein). The coding sequence may be expressed in any or all of the cells of the transgenic animal. For example, the coding sequence may be expressed in the blood, the magnum and/or the sperm of the animal. In a particularly useful embodiment of the invention, the polypeptide is present in an egg, for example, in  
5 the egg white, produced by a transgenic avian.

The present invention also includes methods of dispersing nucleic acid in a cell, for example an avian cell (e.g., an avian embryo cell). These methods include introducing into a cell a nucleic acid and a dispersing agent, for example, a cationic polymer (e.g., polyethylenimine, polylysine, DEAE-dextran, starburst dendrimers  
10 and/or starburst polyamidoamine dendrimers) in an amount that will disperse the nucleic acid in a cell. Typically, the dispersing of the nucleic acid is a homogeneous dispersing. In one embodiment, the nucleic acid includes a transgene. NLS or integrase activity may also be introduced into the cell.

The methods of the invention include introducing the cell into a recipient animal, for example, an avian such as a chicken, wherein the recipient avian produces an offspring which includes the transgene. The cell may be introduced into a recipient animal by any suitable technique.  
15

The present invention also includes the identification of certain regions in the genome which are advantageous for heterologous gene expression. These  
20 regions can be identified by analysis, using methods known in the art, of the transgenic vertebrate animals or cells produced as disclosed herein.

The production of vertebrate animals which are the mature animals developed from the recombinant embryos, ovum and/or sperm of the invention typically are referred to as the G0 generation and are usually hemizygous for each  
25 inserted transgene. The G0 generation may be bred to non-transgenic animals to give rise to G1 transgenic offspring which are also hemizygous for the transgene. The G1 hemizygous offspring may be bred to non-transgenic animals giving rise to G2 hemizygous offspring or may be bred together to give rise to G2 offspring homozygous for the transgene. In one embodiment, hemizygotic G2 offspring  
30 from the same line can be bred to produce G3 offspring homozygous for the transgene. In one embodiment, hemizygous G0 animals are bred together to give rise to homozygous G1 offspring. These are merely examples of certain useful

breeding schemes. The present invention contemplates the employment of any useful breeding scheme such as those known to individuals of ordinary skill in the art.

Any useful combination of features described herein is included within the scope of the present invention provided that the features included in any such combination are not mutually inconsistent as will be apparent from the context, this specification, and the knowledge of one of ordinary skill in the art. For example, the term transgenic can encompass the term transchromosomal and methodologies useful for transgenic animals (e.g., avians) and cells disclosed herein may also be employed for transchromosomal avians and avian cells.

Additional objects and aspects of the present invention will become more apparent upon review of the detailed description set forth below when taken in conjunction with the accompanying figures, which are briefly described as follows.

**15 Brief Description of the Figures**

Fig. 1 illustrates phage integrase-mediated integration. A plasmid vector bearing the transgene includes the attB recognition sequence for the phage integrase. The vector along with integrase-coding mRNA, a vector expressing the integrase, or the integrase protein itself, are delivered into cells or embryos. The integrase recognizes DNA sequences in the avian genome similar to attP sites, termed pseudo-attP, and mediates recombination between the attB and pseudo-attP sites, resulting in the permanent integration of the transgene into the avian genome.

Fig. 2 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells.

Fig. 3 illustrates the results of a puromycin resistance assay to measure phiC31 integrase-mediated integration into chicken cells.

Fig. 4 illustrates phiC31 integrase-mediated integration into quail cells. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase, or a control vector, into QT6 cells, a quail fibrosarcoma cell line. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

Figs. 5A and 5B illustrate that phiC31 integrase can facilitate multiple integrations per avian cell. A puromycin resistance vector bearing an attB site was cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site, and a phiC31 integrase expression vector. After puromycin selection, many puromycin resistant colonies expressed EGFP in all of their cells. Figs. 5A and 5B are the same field of view with EGFP illuminated with ultraviolet light (Fig. 5A) and puromycin resistant colonies photographed in visible light (Fig. 5B). In Fig. 5B, there are 4 puromycin resistant colonies, two of which are juxtaposed at the top. One of these colonies expressed EGFP.

10 Fig. 6 shows maps of the small vectors used for integrase assays.

Fig. 7 shows integrase promotes efficient integration of large transgenes in avian cells.

Fig. 8 shows maps of large vectors used for integrase assays.

15 Fig. 9 illustrates the nucleotide sequence of the integrase-expressing plasmid pCMV-31int (SEQ ID NO: 1).

Fig. 10 illustrates the nucleotide sequence of the plasmid pCMV-luc-attB (SEQ ID NO: 2).

Fig. 11 illustrates the nucleotide sequence of the plasmid pCMV-luc-attP (SEQ ID NO: 3).

20 Fig. 12 illustrates the nucleotide sequence of the plasmid pCMV-pur-attB (SEQ ID NO: 4).

Fig. 13 illustrates the nucleotide sequence of the plasmid pCMV-pur-attP (SEQ ID NO: 5).

25 Fig. 14 illustrates the nucleotide sequence of the plasmid pCMV-EGFP-attB (SEQ ID NO: 6).

Fig. 15 illustrates the nucleotide sequence of the plasmid p12.0-lys-LSIPNMM-CMV-pur-attB (SEQ ID NO: 7).

Fig. 16 illustrates the nucleotide sequence of the plasmid pOMIFN-Ins-CMV-pur-attB (SEQ ID NO: 8).

30 Fig. 17 illustrates the nucleotide sequence of the integrase-expressing plasmid pRSV-Int (SEQ ID NO: 9).

Fig. 18 illustrates the nucleotide sequence of the plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10).

Fig. 19 illustrates the nucleotide sequence of the attP containing polynucleotide SEQ ID NO: 11.

5 Fig. 20 illustrates in schematic from the integration of a heterologous att recombination site into an isolated chromosome. The attB sequence is linked to selectable marker such as a puromycin expression cassette and is flanked by sequences found in the target site of the chromosome to be modified. The DNA is transfected into cells containing the chromosome and stable transfectants are  
10 selected for by drug resistance. Site specific integration may be confirmed by several techniques including PCR.

Fig. 21 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells bearing a wild-type attP sequence.

15 Fig. 22 illustrates the distribution of plasmid DNA in a stage I embryo.

Fig. 23 illustrates the distribution of plasmid DNA in a stage I embryo in the presence of low molecular weight polyethylenimine.

Fig. 24 illustrates the distribution of plasmid DNA in a stage I embryo in the presence of low molecular weight polyethylenimine.

20 Fig. 25 illustrates the integration of a gene of interest (i.e., OMC24-IRES-EPO) into an artificial chromosome by integration (which takes place inside of a host cell) wherein cells containing the recombinant chromosome can be selected for based on hygromycin resistance.

## 25 **Definitions and Abbreviations**

For convenience, definitions of certain terms and certain abbreviations employed in the specification, examples and appended claims are collected here.

Abbreviations used in the present specification include the following: aa, amino acid(s); bp, base pair(s); kb, kilobase(s); mb, megabase(s); att, bacterial  
30 recombination attachment site; IU, infectious units.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates

otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such agents.

The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof.

5   Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

As used herein, an "artificial chromosome" is a nucleic acid molecule that can stably replicate and segregate alongside endogenous chromosomes in a cell. Artificial chromosomes have the capacity to act as gene delivery vehicles by accommodating and expressing foreign genes contained therein. A mammalian artificial chromosome (MAC) refers to chromosomes that have an active mammalian centromere(s). Plant artificial chromosomes, insect artificial chromosomes and avian artificial chromosomes refer to chromosomes that include plant, insect and avian centromeres, respectively. A human artificial chromosome (HAC,) refers to chromosomes that include human centromeres. For exemplary artificial chromosomes, see, for example, U.S. Pat. Nos. 6,025,155, issued February 15, 2000; 6,077,697, issued June 6, 2000; 5,288,625, issued February 22, 2004; 5,712,134, issued January 27, 1998; 5,695,967, issued December 9, 1997; 5,869,294, issued February 9, 1999; 5,891,691, issued April 6, 1999 and 5,721,118, issued February 24, 1998 and published International PCT application Nos, WO 97/40183, published October 30, 1997; and WO 98/08964, published March 5, 1998, the disclosures of which are incorporated herein in their entireties by reference. The term "chromosome" may be used interchangeably with the term "artificial chromosome" as will be apparent based on the context of such use.

Foreign genes that can be contained in artificial chromosome expression systems can include, but are not limited to, nucleic acid that encodes therapeutically effective substances, such as anti-cancer agents, enzymes, hormones and antibodies. Other examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins (reporter genes), such as fluorescent proteins, such as green, blue or red fluorescent proteins (GFP, BFP

and RFP, respectively), other reporter genes, such as beta-galactosidase and proteins that confer drug resistance, such as a gene encoding hygromycin-resistance.

The term "avian" as used herein refers to any species, subspecies or race of organism of the taxonomic class avia, such as, but not limited to chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of Gallus gallus, or chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Australorp, Minorca, Amrox, California Gray), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities. It also includes an individual avian organism in all stages of development, including embryonic and fetal stages. The term "avian" also may denote "pertaining to a bird", such as "an avian (bird) cell."

The terms "chimeric animal" or "mosaic animal" are used herein to refer to an animal in which a nucleotide sequence of interest is found in some but not all cells of the animal, or in which the recombinant nucleic acid is expressed, in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that the recombinant gene is present and/or expressed in some tissues but not others.

The term "coding region" as used herein refers to a continuous linear arrangement of nucleotides which may be translated into a polypeptide. A full length coding region is translated into a full length protein; that is, a complete protein as would be translated in its natural state absent any post-translational modifications. A full length coding region may also include any leader protein sequence or any other region of the protein that may be excised naturally from the translated protein.

The term "cytokine" as used herein refers to any secreted polypeptide that affects a function of cells and modulates an interaction between cells in the immune, inflammatory or hematopoietic response. A cytokine includes, but is not limited to, monokines and lymphokines. Examples of cytokines include, but are not limited to, interferon  $\alpha$ 2b, Interleukin-1 (IL-1), Interleukin-6 (IL-6),

Interleukin-8 (IL-8), Tumor Necrosis Factor- $\alpha$  (TNF- $\alpha$ ) and Tumor Necrosis Factor  $\alpha$  (TNF-  $\alpha$ ).

As used herein, "delivery," which is used interchangeably with "transfection," refers to the process by which exogenous nucleic acid molecules are 5 transferred into a cell such that they are located inside the cell.

As used herein, "DNA" is meant to include all types and sizes of DNA molecules including cDNA, plasmids and DNA including modified nucleotides and nucleotide analogs.

The term "expressed" or "expression" as used herein refers to the 10 transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a region of one of the two nucleic acid strands of the gene. The term "expressed" or "expression" as used herein may also refer to the translation from an RNA molecule to give a protein, a polypeptide or a portion thereof. In one embodiment, for heterologous nucleic acid to be expressed in a 15 host cell, it must initially be delivered into the cell and then, once in the cell, ultimately reside in the nucleus.

The term "gene" or "genes" as used herein refers to nucleic acid sequences 20 that encode genetic information for the synthesis of a whole RNA, a whole protein, or any portion of such whole RNA or whole protein. Genes that are not naturally part of a particular organism's genome are referred to as "foreign genes," "heterologous genes" or "exogenous genes" and genes that are naturally a part of a particular organism's genome are referred to as "endogenous genes". The term "gene product" refers to an RNA or protein that is encoded by the gene. "Endogenous gene products" are RNAs or proteins encoded by endogenous genes. 25 "Heterologous gene products" are RNAs or proteins encoded by "foreign, heterologous or exogenous genes" and are, therefore, not naturally expressed in the cell.

As used herein, the terms "heterologous" and "foreign" with reference to 30 nucleic acids, such as DNA and RNA, are used interchangeably and refer to nucleic acid that does not occur naturally as part of a chromosome, a genome or cell in which it is present or which is found in a location(s) and/or in amounts that differ from the location(s) and/or amounts in which it occurs in nature. It can be nucleic

acid that is not endogenous to the genome, chromosome or cell and has been exogenously introduced into the genome, chromosome or cell. Examples of heterologous DNA include, but are not limited to, DNA that encodes a gene product or gene product(s) of interest, for example, for production of an encoded protein. Examples of heterologous DNA include, but are not limited to, DNA that encodes traceable marker proteins, DNA that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones and as antibodies.

The term "immunoglobulin polypeptide" as used herein refers to a constituent polypeptide of an antibody or a polypeptide derived therefrom. An "immunological polypeptide" may be, but is not limited to, an immunological heavy or light chain and may include a variable region, a diversity region, joining region and a constant region or any combination, variant or truncated form thereof. The term "immunological polypeptides" further includes single-chain antibodies comprised of, but not limited to, an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region and optionally a peptide linker.

The terms "integrase" and "integrase activity" as used herein refer to a nucleic acid recombinase of the serine recombinase family of proteins.

The term "internal ribosome entry sites (IRES)" as used herein refers to a region of a nucleic acid, most typically an RNA molecule, wherein eukaryotic initiation of protein synthesis occurs far downstream of the 5' end of the RNA molecule. A 43S pre-initiation complex comprising the elf2 protein bound to GTP and Met-tRNA<sub>i</sub><sup>Met</sup>, the 40S ribosomal subunit, and factors elf3 and 3lf1A may bind to an "IRES" before locating an AUG start codon. An "IRES" may be used to initiate translation of a second coding region downstream of a first coding region, wherein each coding region is expressed individually, but under the initial control of a single upstream promoter. An "IRES" may be located in a eukaryotic cellular mRNA.

As used herein, the term "large nucleic acid molecules" or "large nucleic acids" refers to a nucleic acid molecule of at least about 0.05 mb in size, greater than 0.5 mb, including nucleic acid molecules at least about 0.6. 0.7, 0.8, 0.9, 1, 5, 10, 30, 50 and 100, 200, 300, 500 mb in size. Large nucleic acid molecules typically can be on the order of about 10 to about 450 or more mb, and can be of

various sizes, such as, for example, from about 250 to about 400 mb, about 150 to about 200 mb, about 90 to about 120 mb, about 60 to about 100 mb and about 15 to 50 mb. A large nucleic acid molecule may be larger than about 8 kb (e.g., about 8 kb to about 1 mb) as will be apparent based on the context.

5 Examples of large nucleic acid molecules include, but are not limited to, natural chromosomes and fragments thereof, especially mammalian chromosomes and fragments thereof which retain a centromere or retain a centromere and telomeres, artificial chromosome expression systems (ACEs which include a mouse centromere; also called satellite DNA-based artificial chromosomes  
10 (SATA Cs); see U.S. Pat. Nos. 6,025,155, issued February 15; and 6,077,697, issued June 20, 2000), mammalian artificial chromosomes (MACs), plant artificial chromosomes, insect artificial chromosomes, avian artificial chromosomes and minichromosomes (see, e.g., U.S. Pat. Nos. 5,712,134, issued January 27, 1998; 5,891,691, issued April 6, 1999; and 5,288,625, issued February 22, 1994). Useful  
15 large nucleic acid molecules can include a single copy of a desired nucleic acid fragment encoding a particular nucleotide sequence, such as a gene of interest, or can carry multiple copies thereof or multiple genes or different heterologous sequences of nucleotides. For example, the chromosomes may carry 1 to about 100 or 1 to about 1000 or even more copies of a gene of interest. Large nucleic acid  
20 molecules can be associated with proteins, for example chromosomal proteins, that typically function to regulate gene expression and/or participate in determining overall structure.

A "nucleic acid fragment of interest" or "nucleotide sequence of interest" may be a trait-producing sequence, by which it is meant a sequence conferring a  
25 non-native trait upon the cell in which the protein encoded by the trait-producing sequence is expressed. The term "non-native" when used in the context of a trait-producing sequence means that the trait produced is different than one would find in an unmodified organism which can mean that the organism produces high amounts of a natural substance in comparison to an unmodified organism, or  
30 produces a non-natural substance. For example, the genome of a bird could be modified to produce proteins not normally produced in birds such as, for example,

useful animal proteins (e.g., human proteins) such as hormones, cytokines and antibodies.

A nucleic acid fragment of interest may additionally be a "marker nucleic acid" or expressed as a "marker polypeptide". Marker genes encode proteins that  
5 can be easily detected in transformed cells and are, therefore, useful in the study of those cells. Examples of suitable marker genes include  $\beta$ -galactosidase, green or yellow fluorescent proteins, enhanced green fluorescent protein, chloramphenicol acetyl transferase, luciferase, and the like. Such regions may also include those 5' noncoding sequences involved with initiation of transcription and translation, such  
10 as the enhancer, TATA box, capping sequence, CAAT sequence, and the like.

As used herein, "nucleic acid" refers to a polynucleotide containing at least two covalently linked nucleotide or nucleotide analog subunits. A nucleic acid can be a deoxyribonucleic acid (DNA), a ribonucleic acid (RNA), or an analog of DNA or RNA. Nucleotide analogs are commercially available and methods of preparing  
15 polynucleotides containing such nucleotide analogs are known (Lin et al. (1994) Nucl. Acids Res. 22:5220-5234; Jellinek et al. (1995) Biochemistry 34:11363-11372; Pagratis et al. (1997) Nature Biotechnol. 15:68-73). The nucleic acid can be single-stranded, double-stranded, or a mixture thereof. For purposes herein, unless specified otherwise, the nucleic acid is double-stranded, or if it is apparent  
20 from the context that the nucleic acid is not double stranded. Nucleic acids include any natural or synthetic linear and sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, certain nucleic acids may be collectively referred to herein as "constructs," "plasmids," or  
25 "vectors."

Techniques useful for isolating and characterizing the nucleic acids and proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals may be consulted to select suitable protocols without undue experimentation. See, for example, Sambrook et  
30 al, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety.

A "nucleoside" is conventionally understood by workers of skill in fields related to the present invention as comprising a monosaccharide linked in glycosidic linkage to a purine or pyrimidine base. A "nucleotide" comprises a nucleoside with at least one phosphate group appended, typically at a 3' or a 5' position (for pentoses) of the saccharide, but may be at other positions of the saccharide. A nucleotide may be abbreviated herein as "nt." Nucleotide residues occupy sequential positions in an oligonucleotide or a polynucleotide. Accordingly a modification or derivative of a nucleotide may occur at any sequential position in an oligonucleotide or a polynucleotide. All modified or derivatized oligonucleotides and polynucleotides are encompassed within the invention and fall within the scope of the claims. Modifications or derivatives can occur in the phosphate group, the monosaccharide or the base.

By way of nonlimiting examples, the following descriptions provide certain modified or derivatized nucleotides. The phosphate group may be modified to a thiophosphate or a phosphonate. The phosphate may also be derivatized to include an additional esterified group to form a triester. The monosaccharide may be modified by being, for example, a pentose or a hexose other than a ribose or a deoxyribose. The monosaccharide may also be modified by substituting hydroxyl groups with hydro or amino groups, by esterifying additional hydroxyl groups. The base may be modified as well. Several modified bases occur naturally in various nucleic acids and other modifications may mimic or resemble such naturally occurring modified bases. Nonlimiting examples of modified or derivatized bases include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-

thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Nucleotides may also be modified to harbor a label. Nucleotides may also bear a fluorescent label or a biotin label.

5       The term “operably linked” refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Control sequences operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The control sequences need not be contiguous with the coding sequence, so long as they function to direct the  
10 expression thereof. For example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered “operably linked” to the coding sequence.

15       “Therapeutic proteins” or “pharmaceutical proteins” include an amino acid sequence which in whole or in part makes up a drug. In one embodiment, a pharmaceutical composition or therapeutic composition includes one or more pharmaceutical proteins or therapeutic proteins.

20       The terms “polynucleotide,” “oligonucleotide,” and “nucleic acid sequence” are used interchangeably herein and include, but are not limited to, coding sequences (polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide in vitro or in vivo when placed under the control of appropriate regulatory or control sequences); control sequences (e.g., translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals, transcription factor binding sites, transcription termination sequences,  
25 upstream and downstream regulatory domains, enhancers, silencers, and the like); and regulatory sequences (DNA sequences to which a transcription factor(s) binds and alters the activity of a gene’s promoter either positively (induction) or negatively (repression)). No limitation as to length or to synthetic origin are suggested by the terms described above.

30       As used herein the terms “peptide,” “polypeptide” and “protein” refer to a polymer of amino acids in a serial array, linked through peptide bonds. A “peptide” typically is a polymer of at least two to about 30 amino acids linked in a

serial array by peptide bonds. The term "polypeptide" includes proteins, protein fragments, protein analogues, oligopeptides and the like. The term "polypeptides" contemplates polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (isolated from an appropriate source such as a bird), or synthesized. The term "polypeptides" further contemplates polypeptides as defined above that include chemically modified amino acids or amino acids covalently or noncovalently linked to labeling moieties.

The terms "percent sequence identity" or "percent sequence similarity" as used herein refer to the degree of sequence identity between two nucleic acid sequences or two amino acid sequences as determined using the algorithm of Karlin & Atschul, Proc. Natl. Acad. Sci. 87: 2264-2268 (1990), modified as in Karlin & Atschul, Proc. Natl. Acad. Sci. 90: 5873-5877 (1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Atschul et al, 1990, T. Mol. Biol. 215: 403-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, word length = 12, to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, word length = 3, to obtain amino acid sequences homologous to a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Atschul et al, Nucl. Acids Res. 25: 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. Other algorithms, programs and default settings may also be suitable such as, but not only, the GCG-Sequence Analysis Package of the U.K. Human Genome Mapping Project Resource Centre that includes programs for nucleotide or amino acid sequence comparisons. Examples of useful algorithms are FASTA and BESTFIT.

The term "promoter" as used herein refers to the DNA sequence that determines the site of transcription initiation by an RNA polymerase. A "promoter-proximal element" is a regulatory sequence generally within about 200 base pairs of the transcription start site.

The term "pseudo-recombination site" as used herein refers to a site at which an integrase can facilitate recombination even though the site may not have a

sequence identical to the sequence of its wild-type recombination site. For example, a phiC31 integrase and vector carrying a phiC31 wild-type recombination site can be placed into an avian cell. The wild-type recombination sequence aligns itself with a sequence in the avian cell genome and the integrase facilitates a 5 recombination event. When the sequence from the genomic site in the avian cell, where the integration of the vector took place, is examined, the sequence at the genomic site typically has some identity to, but may not be identical with, the wild-type bacterial genome recombination site. The recombination site in the avian cell genome is considered to be a pseudo-recombination site (e.g., a pseudo-attP site) at 10 least because the avian cell is heterologous to the normal phiC31 phage/bacterial cell system. The size of the pseudo-recombination site can be determined through the use of a variety of methods including, but not limited to, (i) sequence alignment comparisons, (ii) secondary structural comparisons, (iii) deletion or point mutation analysis to find the functional limits of the pseudo-recombination site, and (iv) 15 combinations of the foregoing.

The terms "recombinant cell" and "genetically transformed cell" refer to a cell comprising a combination of nucleic acid segments not found in a single cell with each other in nature. A new combination of nucleic acid segments can be introduced into an organism using a wide array of nucleic acid manipulation 20 techniques available to those skilled in the art. The recombinant cell may harbor a vector that is extragenomic, i.e. that does not covalently insert into the cellular genome, including a non-nuclear (e.g. mitochondrial) genome(s). A recombinant cell may further harbor a vector or a portion thereof that is intragenomic, i.e. covalently incorporated within the genome of the recombinant cell.

25 The term "recombination site" as used herein refers to a polynucleotide stretch comprising a recombination site normally recognized and used by an integrase. For example,  $\lambda$  phage is a temperate bacteriophage that infects *E. coli*. The phage has one attachment site for recombination (attP) and the *E. coli* bacterial genome has an attachment site for recombination (attB). Both of these sites are 30 recombination sites for  $\lambda$  integrase. Recombination sites recognized by a particular integrase can be derived from a homologous system and associated with

heterologous sequences, for example, the attP site can be placed in other systems to act as a substrate for the integrase.

The terms "recombinant nucleic acid" and "recombinant DNA" as used herein refer to combinations of at least two nucleic acid sequences that are not naturally found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques. A recombinant polypeptide may be distinct from a naturally occurring polypeptide either in its location, purity or structure. Generally, a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

As used herein, the term "satellite DNA-based artificial chromosome (SATAC)" (e.g., ACE) is a type of artificial chromosome. These artificial chromosomes are substantially all neutral non-coding sequences (heterochromatin) except for foreign heterologous, typically gene-encoding nucleic acid, that is present within (see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697, issued June 20, 2000 and International PCT application No. WO 97/40183, published October 30, 1997).

The term "source of integrase activity" as used herein refers to a polypeptide or multimeric protein having serine recombinase (integrase) activity in an avian cell. The term may further refer to a polynucleotide encoding the serine recombinase, such as an mRNA, an expression vector, a gene or isolated gene that may be expressed as the recombinase-specific polypeptide or protein.

As used herein the term "therapeutic substance" refers to a component that comprises a substance which can provide for a therapeutic effect, for example, a therapeutic protein.

"Transchromosomal avian" means an avian which contains an artificial chromosome in some or all of its cells. A transchromosomal avian can include the artificial chromosome in its germ cells.

The term "transcription regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the transcriptional expression of the gene. Exemplary transcription regulatory sequences include enhancer elements, hormone response elements, 5 steroid response elements, negative regulatory elements, and the like.

The term "transfection" as used herein refers to the process of inserting a nucleic acid into a host cell. Many techniques are well known to those skilled in the art to facilitate transfection of a nucleic acid into an eukaryotic cell. These methods include, for instance, treating the cells with high concentrations of salt 10 such as a calcium or magnesium salt, an electric field, detergent, or liposome mediated transfection, to render the host cell competent for the uptake of the nucleic acid molecules, and by such methods as micro-injection into a pro-nucleus, sperm-mediated and restriction-mediated integration.

The term "transformed" as used herein refers to a heritable alteration in a 15 cell resulting from the uptake of a heterologous DNA.

As used herein, the term "transgene" means a nucleic acid sequence that is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is 20 inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout).

As used herein, a "transgenic avian" is any avian, as defined herein, in which one or more of the cells of the avian contain heterologous nucleic acid 25 introduced by manipulation, such as by transgenic techniques. The nucleic acid may be introduced into a cell, directly or indirectly, by introduction into a precursor of the cell by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. Genetic manipulation also includes classical cross-breeding, or in vitro fertilization. A recombinant DNA molecule 30 may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The term “trisomic” as used herein refers to a cell or animal, such as an avian cell or bird that has a  $2n+1$  chromosomal complement, where  $n$  is the haploid number of chromosomes, for the animal species concerned.

5 The terms “vector” or “nucleic acid vector” as used herein refer to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule (RNA or DNA) that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. The term “expression vector” as used herein refers to a nucleic acid vector that comprises a transcription regulatory region operably linked to a site wherein is, or can be, inserted, a nucleotide sequence to be transcribed and, optionally, to be expressed, for instance, but not limited to, a sequence coding at least one polypeptide.

10

### **Detailed Description**

15 The present invention provides for recombinant vertebrate cells (e.g., transgenic or transchromosomal avian cells) and transgenic vertebrate animals (e.g., transgenic or transchromosomal avians) and methods of making the cells and the animals. For example, the invention provides for methods of inserting nucleotide sequences into the genome of vertebrate animals or into the cells of vertebrate animals in a site specific manner. Examples of vertebrates include, 20 without limitation, birds, mammals, fish, reptiles and amphibians. Examples of mammals include sheep, goats and cows. In one certain embodiment of the invention, the vertebrate animals are birds or avians. Examples of birds include, without limitation, chickens, turkeys, ducks, geese, quail, pheasants, parrots, finches, hawks, crows and ratites including ostriches, emu and cassowary.

25 In one embodiment, the present invention provides for methods of inserting nucleotide sequences into the genome of an animal using methods of transgenesis based on site specific integration, for example, site specific integrase mediated-transgenesis. The present invention contemplates any useful method of integrase mediated transgenesis including but not limited to, transgenesis mediated by serine 30 recombinases and tyrosine recombinases. Serine recombinases are well known in the art and include without limitation, EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1,  $\Phi$ 370.1,

Φ105, ΦFC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, ΦRv1, Y4bA, Bja, SsoISC1904b, SsoISC1904a, Aam, MjaMJ1004, Pab, SsoISC1913, HpyIS607, McrV0921, MtuRv0921, MtuRv2979c, MtuRv2792c, MtuISY349, MtuRv3828c, SauSK1, 5 Spy, EcoTn21, Mlo92, EcoTn3, Lla, Cpe, SauSK41, BmeTn5083, SfaTn917, Bme53, Ran, RmzY4CG, SarpNL1, Pje, Xan, ISXc5, Pae, Xca, Req, Mlo90, PpsTn5501, pMER05, Cgl, MuGin, StyHin, Xfa911, Xfa910, Rrh, SauTn552 and Aac serine recombinases. Tyrosine recombinases well known in the art include without limitation, BS codV, BS ripX, BS ydcL, CB tnpA, Col1D, CP4, Cre, D29, 10 DLP12, DN int, EC FimB, EC FimE, EC orf, EC xerC, EC xerD, Φ11, Φ13, Φ80, Φadh, ΦCTX, ΦLC3, FLP, ΦR73, HIorf, HI rci, HI xerC, HI xerD, HK22, HP1, L2, L5, L54, λ, LL orf, LL xerC, LO L5, MJ orf, ML orf, MP int, MT int, MT orf, MV4, P186, P2, P21, P22, P4, P434, PA sss, PM fimB, pAE1, pCL1, pKD1, pMEA, pSAM2, pSB2, pSB3, pSDL2, pSE101, pSE211, pSM1, pSR1, pWS58, 15 R721, Rci, SF6, SLP1, SM orf, SsrA, SSV1, T12, Tn21, Tn4430, Tn554a, Tn554b, Tn7, Tn916, Tuc, WZ int, XisA and XisC. Other enzymes which may be useful for mediation of transgenesis in accordance with the present invention include, certain transposases, invertases and resolvases.

In certain instances, integration host factors (IHF) may be necessary for the 20 integration of nucleotide sequences of the invention into the genome of cells as disclosed herein. In such a case, the integration host factors may be delivered to the cells directly or they may be delivered to the cells in the form of a nucleic acid which, in the case of RNA, is translated to produce the IHF or, in the case of DNA, is transcribed and translated to produce the IHF.

The present invention contemplates the use of any system capable of site 25 specifically inserting a nucleotide sequence of interest into the genome of a cell, for example, to produce a transgenic vertebrate animal. Typically, although not exclusively, these systems require at least three components: 1) a sequence in the genome which specifies the site of insertion; 2) a nucleotide sequence which is directed to the site of insertion and an enzyme which catalyzes the insertion of the 30 nucleotide sequence into the genome at the site of insertion. Many enzymes, including integrases, which are capable of site specifically inserting nucleotide

sequences into the genome have been characterized. Examples of these enzymes are disclosed in for example, Esposito et al (1997) Nucleic Acids Research, 25;3605-3614 and Nunes-Düby et al (1998) Nucleic Acids Research, 26; 391-406. The disclosure of each of these references is incorporated herein in their entirety.

5 In one embodiment of the present invention, a serine recombinase is employed. Serine recombinase integrase mediates recombination between an attB site on a transgene vector and an attP or a pseudo attP site on a chromosome. In the method of the invention for integrase-mediated transgenesis, a heterologous wild-type attP site can be integrated into a nuclear genome to create a transgenic  
10 cell line or a transgenic vertebrate animal, such as an avian. A serine recombinase (integrase) and an attB-bearing transgene vector are then introduced into cells harboring the heterologous attP site, or into embryos derived from animals which bear the attP recombination site. The locations of attP and attB may be reversed such that the attB site is inserted into a chromosome and the attP sequence resides  
15 in an incoming transgene vector. In either case, the att site of the introduced vector would then preferentially recombine with the integrated heterologous att site in the genome of the recipient cell.

The methods of the invention are based, in part, on the discovery that there exists in vertebrate animal genomes, such as avian genomes, a number of specific  
20 nucleic acid sequences, termed pseudo-recombination sites, the sequences of which may be distinct from wild-type recombination sites but which can be recognized by a site-specific integrase and used to promote the efficient insertion of heterologous genes or polynucleotides into the targeted nuclear genome. The inventors have identified pseudo-recombination sites in avian cells capable of recombining with a  
25 recombination site, such as an attB site within a recombinant nucleic acid molecule introduced into the target avian cell. The invention is also based on the prior integration of a heterologous att recombination site, typically isolated from a bacteriophage or a modification thereof, into the genome of the target avian cell.

Integration into a predicted chromosomal site is useful to improve the  
30 predictability of expression, which is particularly advantageous when creating transgenic avians. Transgenesis by methods that result in insertion of the transgene

into random positions of the avian genome is unpredictable since the transgene may not express at the expected levels or in the predicted tissues.

The invention as disclosed herein, therefore, provides methods for site-specifically genetically transforming an avian nuclear genome. In general, an avian cell having a first recombination site in the nuclear genome is transformed with a site-specific polynucleotide construct comprising a second recombination sequence and one or more polynucleotides of interest. Into the same cell, integrase activity may be introduced that specifically recognizes the first and second recombination sites under conditions such that the polynucleotide sequence of interest is inserted into the nuclear genome via an integrase-mediated recombination event between the first and second recombination sites.

The integrase activity, or a source thereof, can be introduced into the cell prior to, or concurrent with, the introduction of the site-specific construct. The integrase can be delivered to a cell as a polypeptide, or by expressing the integrase from a source polynucleotide such as an mRNA or from an expression vector that encodes the integrase, either of which can be delivered to the target cell before, during or after delivery of the polynucleotide of interest. Any integrase that has activity in a cell may be useful in the present invention, including HK022 (Kolot et al, Biotechnol. Bioeng., 84: 56-60 (2003)). In one embodiment, the integrase is a serine recombinase as described, for example, by Smith & Thorpe, in Mol. Microbiol., 44: 299-307 (2002). For example, the integrase may be TP901-1 (Stoll et al, J. Bact., 184: 3657-3663 (2002); Olivares et al, Gene, 278:167-176 (2001) or the integrase from the phage phiC31.

The nucleotide sequence of the junctions between an integrated transgene into the attP (or attB site) would be known. Thus, a PCR assay can be designed by one of skill in the art to detect when the integration event has occurred. The PCR assay for integration into a heterologous wild-type attB or attP site can also be readily incorporated into a quantitative PCR assay using TAQMAN™ or related technology so that the efficiency of integration can be measured.

In one embodiment, the minimal attB and attP sites able to catalyze recombination mediated by the phiC31 integrase are 34 and 39 bp, respectively. In cell lines that harbor a heterologous integrated attP site, however, integrase may

have a preference for the inserted attP over any pseudo-attP sites of similar length, because pseudo-attP sites have very low sequence identity (for example, between 10 to 50% identity) compared to the more efficient wild-type attP sequence. It is within the scope of the methods of the invention, however, for the recombination site within the target genome to be a pseudo-att site such as a pseudo-attP site or an attP introduced into a genome.

The sites used for recognition and recombination of phage and bacterial DNAs (the native host system) are generally non-identical, although they typically have a common core region of nucleic acids. In one embodiment, the bacterial sequence is called the attB sequence (bacterial attachment) and the phage sequence is called the attP sequence (phage attachment). Because they are different sequences, recombination can result in a stretch of nucleic acids (for example, attL or attR for left and right) that is neither an attB sequence or an attP sequence, and likely is functionally unrecognizable as a recombination site to the relevant enzyme, thus removing the possibility that the enzyme will catalyze a second recombination reaction that would reverse the first.

The integrase may recognize a recombination site where sequence of the 5' region of the recombination site can differ from the sequence of the 3' region of the recombination sequence. For example, for the phage phiC31 attP (the phage attachment site), the core region is 5'-TTG-3' the flanking sequences on either side are represented here as attP5' and attP3', the structure of the attP recombination site is, accordingly, attP5'-TTG-attP3'. Correspondingly, for the native bacterial genomic target site (attB) the core region is 5'-TTG-3', and the flanking sequences on either side are represented here as attB5' and attB3', the structure of the attB recombination site is, accordingly, attB5'-TTG-attB3'. After a single-site, phiC31 integrase-mediated recombination event takes place between the phiC31 phage and the bacterial genome, the result is the following recombination product: attB5'-TTG-attP3'{phiC31 vector sequences}-attP5'-TTG-attB3'. In the method of invention, the attB site will be within a recombinant nucleic acid molecule that may be delivered to a target cell. The corresponding attP (or pseudo-attP) site will be within the cell nuclear genome. Consequently, after phiC31 integrase mediated recombination, the recombination product, the nuclear genome with the integrated

heterologous polynucleotide will have the sequence attP5'-TTG-attB3'{heterologous polynucleotide}-attB5'-TTG-attP3'. Typically, after recombination the post-recombination recombination sites are no longer able to act as substrate for the phiC31 integrase. This results in stable integration with little or  
5 no integrase mediated excision.

While the one useful recombination site to be included in the recombinant nucleic acid molecules and modified chromosomes of the present invention is the attP site, it is contemplated that any attP-like site may be used if compatible with the attB site. For instance, any pseudo-attP site of the chicken genome may be  
10 identified according to the methods of Example 7 herein and used as a heterologous att recombination site. For example, such attP-like sites may have a sequence that is greater than at least 25% identical to SEQ ID NO: 11 as shown in Fig. 19, such as described in Groth et al, Proc. Natl. Acad. Sci. U.S.A. 97: 5995-6000 (2000) incorporated herein by reference in its entirety. In one embodiment,  
15 the selected site will have a similar degree of efficiency of recombination, for example, at least the same degree of efficiency of recombination as the attP site (SEQ ID NO: 11) itself.

In the present invention, the recipient cell population may be an isolated cell line such as, for example, DF-1 chicken fibroblasts, chicken DT40 cells or a  
20 cell population derived from an early stage embryo, such as a chicken stage I embryo or mid stage or late stage (e.g., stage X) embryos. One useful avian cell population is blastodermal cells isolated from a stage X avian embryo. The methods of the present invention, therefore, include steps for the isolation of blastodermal cells that are then suspended in a cell culture medium or buffer for  
25 maintaining the cells in a viable state, and which allows the cell suspension to contact the nucleic acids of the present invention. It is also within the scope of the invention for the nucleic acid construct and the source of integrase activity to be delivered directly to an avian embryo such as a blastodermal layer, or to a tissue layer of an adult bird such as the lining of an oviduct.

30 When the recipient cell population is isolated from an early stage avian embryo, the embryos must first be isolated. For stage I avian embryos from, for example, a chicken, a fertilized ovum is surgically removed from a bird before the

deposition of the outer hard shell has occurred. The nucleic acids for integrating a heterologous nucleic acid into a recipient cell genome may then be delivered to isolated embryos by lipofection, microinjection (as described in Example 6 below) or electroporation and the like. After delivery of the nucleic acid, the transfected 5 embryo and its yolk may be deposited into the infundibulum of a recipient hen for the deposition of egg white proteins and a hard shell, and laying of the egg. Stage X avian embryos are obtained from freshly laid fertilized eggs and the blastodermal cells isolated as a suspension of cells in a medium, as described in Example 4 below. Isolated stage X blastodermal cell populations, once transfected, may be 10 injected into recipient stage X embryos and the hard shell eggs resealed according to the methods described in U.S. Patent No. 6,397,777, issued June 4, 2002, the disclosure of which is incorporated in its entirety by reference herein.

In one embodiment of the invention, once a heterologous nucleic acid is delivered to the recipient cell, integrase activity is expressed. The expressed 15 integrase (or injected integrase polypeptide) then mediates recombination between the att site of the heterologous nucleic acid molecule, and the att (or pseudo att) site within the genomic DNA of the recipient avian cell.

It is within the scope of the present invention for the integrase-encoding sequence and a promoter operably linked thereto to be included in the delivered 20 nucleic acid molecule and that expression of the integrase activity occurs before integration of the heterologous nucleic acid into the cell genome. In one embodiment, an integrase-encoding nucleic acid sequence and associated promoter are in an expression vector that may be co-delivered to the recipient cell with the heterologous nucleic acid molecule to be integrated into the recipient genome.

25 One suitable integrase expressing expression vector for use in the present invention is pCMV-C31int (SEQ ID NO: 1) as shown in Fig. 9, and described in Groth et al, Proc. Natl. Acad. Sci. U.S.A. 97: 5995-6000 (2000), incorporated herein by reference in its entirety. In pCMV-C31int, expression of the integrase-encoding sequence is driven by the CMV promoter. However, any promoter may 30 be used that will give expression of the integrase in a recipient cell, including operably linked avian-specific gene expression control regions of the avian

ovalbumin, lysozyme, ovomucin, ovomucoid gene loci, viral gene promoters, inducible promoters, the RSV promoter and the like.

- The recombinant nucleic acid molecules of the present invention for delivery of a heterologous polynucleotide to the genome of a recipient cell may 5 comprise a nucleotide sequence encoding the attB attachment site of *Streptomyces ambofaciens* as described in Thorpe & Smith, Proc. Natl. Acad. Sci. U.S.A. 95: 5505-5510 (1998). The nucleic acid molecule of the present invention may further comprise an expression cassette for the expression in a recipient cell of a heterologous nucleic acid encoding a desired heterologous polypeptide.
- 10 Optionally, the nucleic acid molecules may also comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP, and the like.

It is contemplated that the expression cassette, for introducing a desired heterologous polypeptide, comprises a promoter operably linked to a nucleic acid encoding the desired polypeptide and, optionally, a polyadenylation signal 15 sequence. Exemplary nucleic acids suitable for use in the present invention are more fully described in the examples below.

In one embodiment of the present invention, following delivery of the nucleic acid molecule and a source of integrase activity into a cell population, for example, an avian cell population, the cells are maintained under culture conditions 20 suitable for the expression of the integrase and/or for the integrase to mediate recombination between the recombination site of the nucleic acid and recombination site in the genome of a recipient cell. When the recipient cell is cultured in vitro, such cells may be incubated at 37° Celsius. For example, chicken early stage blastodermal cells may be incubated at 37° Celsius. They may then be 25 injected into an embryo within a hard shell, which is resealed for incubation until hatching. Alternatively, the transfected cells may be maintained in in vitro culture.

In one embodiment, the present invention provides methods for the site-specific insertion of a heterologous nucleic acid molecule into the nuclear genome of a cell by delivering to a target cell that has a recombination site in its nuclear 30 genome, a source of integrase activity, a site-specific construct that has another recombination site and a polynucleotide of interest, and allowing the integrase activity to facilitate a recombination event between the two recombination sites,

thereby integrating the polynucleotide of interest into the nuclear genome.

(a) Expression vector nucleic acid molecules: A variety of recombinant nucleic acid expression vectors are suitable for use in the practice of the present invention. The site-specific constructs described herein can be constructed utilizing methodologies well known in the art of molecular biology (see, for example, Ausubel or Maniatis) in view of the teachings of the specification. As described above, the constructs are assembled by inserting into a suitable vector backbone a recombination site such as an attP or an attB site, a polynucleotide of interest operably linked to a gene expression control region of interest and, optionally a sequence encoding a positive selection marker. Polynucleotides of interest can include, but are not limited to, expression cassettes encoding a polypeptide to be expressed in the transformed cell or in a transgenic vertebrate animal derived therefrom. The site-specific constructs are typically, though not exclusively, circular and may also contain selectable markers, an origin of replication, and other elements.

Any of the vectors of the present invention may also optionally include a sequence encoding a signal peptide that directs secretion of the polypeptide expressed by the vector from the transgenic cells, for instance, from tubular gland cells of the oviduct of an avian. In one embodiment, this aspect of the invention effectively broadens the spectrum of exogenous proteins that may be deposited in the whites of avian eggs using the methods of the invention. Where an exogenous polypeptide would not otherwise be secreted, the vector bearing the coding sequence can be modified to comprise, for instance, about 60 bp encoding a signal peptide. The DNA sequence encoding the signal peptide may be inserted in the vector such that the signal peptide is located at the N-terminus of the polypeptide encoded by the vector.

The expression vectors of the present invention can comprise a transcriptional regulatory region, for example, an avian transcriptional regulatory region, for directing expression of either fusion or non-fusion proteins. With fusion vectors, a number of amino acids are usually added to the desired expressed target gene sequence such as, but not limited to, a polypeptide sequence for thioredoxin. A proteolytic cleavage site may further be introduced at a site

between the target recombinant protein and the fusion sequence. Additionally, a region of amino acids such as a polymeric histidine region may be introduced to allow binding of the fusion protein to metallic ions such as nickel bonded to a solid support, for purification of the fusion protein. Once the fusion protein has been purified, the cleavage site allows the target recombinant protein to be separated from the fusion sequence. Enzymes suitable for use in cleaving the proteolytic cleavage site include, but are not limited to, Factor Xa and thrombin. Fusion expression vectors that may be useful in the present invention include pGex (Amrad Corp., Melbourne, Australia), pRIT5 (Pharmacia, Piscataway, NJ) and pMAL (New England Biolabs, Beverly, MA), that fuse glutathione S-transferase, protein A, or maltose E binding protein, respectively, to a desired target recombinant protein.

Epitope tags are short peptide sequences that are recognized by epitope specific antibodies. A fusion protein comprising a recombinant protein and an epitope tag can be simply and easily purified using an antibody bound to a chromatography resin, for example. The presence of the epitope tag furthermore allows the recombinant protein to be detected in subsequent assays, such as Western blots, without having to produce an antibody specific for the recombinant protein itself. Examples of commonly used epitope tags include V5, glutathione-S-transferase (GST), hemagglutinin (HA), the peptide Phe-His-His-Thr-Thr, chitin binding domain, and the like.

Exemplary gene expression control regions for use in cells such as avian cells (e.g., chicken cells) include, but are not limited to, avian specific promoters such as the chicken lysozyme, ovalbumin, or ovomucoid promoters, and the like. Particularly useful in avian systems are tissue-specific promoters such as avian oviduct promoters that allow for expression and delivery of a heterologous polypeptide to an egg white.

Viral promoters serve the same function as bacterial or eukaryotic promoters and either provide a specific RNA polymerase in trans (bacteriophage T7) or recruit cellular factors and RNA polymerase (SV40, RSV, CMV). Viral promoters can be useful as they are generally particularly strong promoters. One useful promoter for employment in avian cells is the RSV promoter.

Selection markers are valuable elements in expression vectors as they provide a means to select for growth of only those cells that contain a vector. Common selectable marker genes include those for resistance to antibiotics such as ampicillin, puromycin, tetracycline, kanamycin, bleomycin, streptomycin, 5 hygromycin, neomycin, ZEOCINTM, and the like.

Another element useful in an expression vector is an origin of replication. Replication origins are unique DNA segments that contain multiple short repeated sequences that are recognized by multimeric origin-binding proteins and that play a key role in assembling DNA replication enzymes at the origin site. Suitable origins 10 of replication for use in expression vectors employed herein include E. coli oriC, colE1 plasmid origin, and the like.

A further useful element in an expression vector is a multiple cloning site or polylinker. Synthetic DNA encoding a series of restriction endonuclease recognition sites is inserted into a vector, for example, downstream of the promoter 15 element. These sites are engineered for convenient cloning of DNA into the vector at a specific position.

Elements such as the foregoing can be combined to produce expression vectors suitable for use in the methods of the invention. Those of skill in the art will be able to select and combine the elements suitable for use in their particular 20 system in view of the teachings of the present specification.

Provided for is the stable introduction of a large DNA molecule into the cell of an avian. In one particularly useful embodiment, the large DNA molecule is a chromosome. The chromosomes to be introduced into cells of an avian may be referred to herein as "artificial chromosomes"; however, the term "artificial 25 chromosome" is not a limiting term and any useful large DNA molecule or chromosome may be employed in the present invention.

The present invention provides modified chromosomes, which are either isolated chromosomes or artificial chromosomes, which function as useful vectors to shuttle transgenes or gene clusters into the genome. By delivering the modified 30 or artificial chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic or transchromosomal. Typically, an additional or trisomic chromosome will not affect the subsequent development of the recipient

cell and/or an embryo, nor interfere with the reproductive capacity of an adult developed from such cells or embryos. The chromosome also should be stable within chicken cells. An effective method is also required to isolate a population of chromosomes for delivery into chicken embryos or early cells.

5 Chickens that are trisomic for microchromosome 16 have been described (Miller et al, Proc. Natl. Acad. Sci. U.S.A. 93: 3958-3962 (1996); Muscarella et al, J. Cell Biol. 101: 1749-1756 (1985). In these cases, triploidy and trisomy occurred naturally, and illustrate that an extra copy of one or more of the chicken chromosomes is compatible with normal development and reproductive capacity.

10 The transchromosomal avians resulting from the cellular introduction of an artificial chromosome typically will comprise cells which include the normal complement of chromosomes plus at least one additional chromosome. In one embodiment, about 0.001% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 0.1% to 100% of the cells  
15 of the avian will include an additional chromosome. In another embodiment, about 5% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 10% to 100% of the cells of the avian will include an additional chromosome. In another embodiment, about 50% to 100% of the cells of the avian will include an additional chromosome. In one particularly useful  
20 embodiment, the additional chromosome is transmitted through the germ-line of the transchromosomal avian and many, for example, most (i.e., more than 50%) of the cells of the offspring avians will include the additional chromosome. The invention contemplates the introduction and propagation of any useful number of chromosomes into the cell(s) of a transgenic avian or isolated avian cells. For  
25 example, the invention contemplates one artificial chromosome or two artificial chromosomes or three artificial chromosomes stably incorporated into the genome of the cell(s) of a transchromosomal avian or isolated avian cells.

Any or all tissues of the transchromosomal avian can include the artificial chromosome. In one useful embodiment, one or more cells of the oviduct of the  
30 avians include the additional chromosome. For example, tubular gland cells of the oviduct may include the additional chromosome.

A number of artificial chromosomes are useful in the methods of the

invention, including, for instance, a human chromosome modified to work as an artificial chromosome in a heterologous species as described, for example, for mice (Tomizuka et al, Proc. Natl. Acad. Sci. U.S.A. 97: 722-727 (2000); for cattle (Kuroiwa et al, Nat. Biotechnol. 20: 889-894 (2002); a mammalian artificial 5 chromosome used in mice (Co et al, Chromosome Res. 8: 183-191 (2000).

Examples of large nucleic acid molecules include, but are not limited to, natural chromosomes and fragments thereof, for example, chromosomes (e.g., mammalian chromosomes) and fragments thereof which retain a centromere, artificial chromosome expression systems (satellite DNA-based artificial 10 chromosomes (SATAcs); see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697 issued June 20, 2000, the disclosures of which are incorporated herein in their entirety by reference), mammalian artificial chromosomes (MACs) (e.g., HACs), plant artificial chromosomes, insect artificial chromosomes, avian 15 artificial chromosomes and minichromosomes (see, e.g., U.S. Pat. Nos. 5,712,134 issued January 27, 1998; 5,891,691, issued April 6, 1999; 5,288,625, issued February 22, 1994; 6,743,967 issued June 1, 2004; and U.S. Patent Application Nos. 10/235,119, published June 19, 2003, the disclosure of each of these six 20 patents and the patent application are incorporated herein in their entirety by reference). Also contemplated for use herein are YACs, BACs, bacteriophage-derived artificial chromosomes (BBPACs), cosmid or P1 derived artificial chromosomes (PACs).

As used herein, a large nucleic acid molecule such as artificial chromosomes can stably replicate and segregate alongside endogenous chromosomes in a cell. It has the capacity to act as a gene delivery vehicle by 25 accommodating and expressing foreign genes contained therein. A mammalian artificial chromosome (MAC) refers to chromosomes that have an active mammalian centromere(s). Plant artificial chromosomes, insect artificial chromosomes and avian artificial chromosomes refer to chromosomes that include plant, insect and avian centromeres, respectively. A human artificial chromosome 30 (HAC,) refers to chromosomes that include human centromeres. For exemplary artificial chromosomes, see, e.g., U.S. Pat. Nos. 6,025,155, issued February 15, 2000; 6,077,697, issued June 20, 2000; 5,288,625, issued February 22, 1994;

5,712,134, issued January 27, 1998; 5,695,967, issued December 9, 1997; 5,869,294, issued February 9, 1999; 5,891,691, issued April 6, 1999 and 5,721,118, issued February 24, 1998 and published International PCT application Nos., WO 97/40183, published October 30, 1997 and WO 98/08964, published March 5, 5 1998, the disclosure of each of these eight patents and two PCT applications are incorporated in their entirety herein by reference.

The large nucleic acid molecules (e.g., chromosomes) can include a single copy of a desired nucleic acid fragment encoding a particular nucleotide sequence, such as a gene of interest, or can carry multiple copies thereof or multiple genes, 10 different heterologous nucleotide sequences or expression cassettes or may encode one or more heterologous transcripts each encoding more than one useful protein product (for example, the transcript(s) may comprise an IRES). Any useful IRES may be employed in the invention. See, for example, US Patent No. 4,937,190, issued January 26, 1990; Nature (1988) 334:320-325; J Virol (1988) 62:3068- 15 3072; Cell (1992) 68:119-131; J Virol (1990) 64:4625-4631; and J Virol (1992) 66:1476-1483, the disclosures of which are incorporated in their entirety herein by reference, which disclose useful IRESs. For example, the nucleic acid molecules can carry 40 or even more copies of genes of interest. The large nucleic acid molecules can be associated with proteins, for example, chromosomal proteins, 20 that typically function to regulate gene expression and/or participate in determining overall structure (e.g., nucleosomes).

Certain useful artificial chromosomes, such as satellite DNA-based artificial chromosomes, can include substantially all neutral non-coding sequences (heterochromatin) except for foreign heterologous, typically gene-encoding, nucleic acid (see U.S. Pat. Nos. 6,025,155, issued February 15, 2000 and 6,077,697, issued 25 June 20, 2000 and International PCT application No. WO 97/40183, published October 30, 1997 and Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172, the disclosures of these two patents, the PCT application and the publication are incorporated in their entirety herein by reference). Foreign genes (i.e., 30 nucleotide sequences of interest) contained in these artificial chromosomes can include, but are not limited to, nucleic acid that encodes therapeutically effective substances (e.g., therapeutic proteins such as those disclosed elsewhere herein and

traceable marker proteins (reporter genes), such as fluorescent proteins, such as green, blue or red fluorescent proteins (GFP, BFP and RFP, respectively), other reporter genes, such as beta-galactosidase and proteins that confer drug resistance, such as a gene encoding hygromycin-resistance.

5        In one useful embodiment, the artificial chromosomes employed herein do not interfere with the host cells' processes and can be easily purified by useful purification methods such as large-scale by high-speed flow cytometry (see, for example, de Jong, G, et al. Cytometry 35: 129-33, 1999). Such artificial chromosomes are useful for the production of transchromosomal chickens  
10 produced by introduction of the chromosomes into certain cells, for example, the germline cells, of an avian. In one particularly useful embodiment of the present invention, the transchromosomal chickens are produced by microinjection of the chromosomes, for example, cytoplasmic injection of the chromosomes into avian embryos, for example, early stage embryos such as a Stage I embryos, see, for  
15 example, US Patent Application No. 10/679,034, filed October 2, 2003, the disclosure of which is incorporated in its entirety herein by reference.

In one embodiment, heterologous nucleic acid is introduced into an artificial chromosome. Any useful method to introduce the nucleic acid into the chromosome may be employed in the invention. Thereafter, the artificial  
20 chromosomes are isolated in a mixture substantially free of other chromosomes or cellular material. For example, artificial chromosomes may be isolated by flow cytometry (e.g., dual laser high-speed flow cytometer as described previously (de Jong, G, et al. Cytometry 35: 129-33, 1999). See, for example, US Patent Application Publication No. 20030113917, published June 19, 2003, the disclosure  
25 of which is incorporated in its entirety herein by reference.

In accordance with the present invention, any useful number of artificial chromosomes may be introduced into an avian cell (e.g., injected), for example, an avian germinal cell such as a cell of an ova, an embryo or a germinal disc of an avian egg. Any useful method of introducing the chromosomes into the avian cell  
30 is contemplated for use in the present invention. In addition, the invention contemplates the introduction of any useful number of chromosomes into an avian cell. For example, and without limitation, the invention contemplates the

introduction of 1 to about 1,000,000 chromosomes injected per egg. In one embodiment, 1 to about 100,000 chromosomes are injected per egg. In another embodiment about 5 to about 100,000 artificial chromosomes are injected per egg. For example, about 10 to about 50,000 chromosomes may be injected per egg.

5       In one embodiment, there is a lower hatch rate for eggs injected with more than a certain number of chromosomes. In one embodiment, an injection of over 100,000 chromosomes reduces or brings the hatch rate to zero. In another embodiment, an injection of over 20,000 chromosomes reduces or brings the hatch rate to zero. In another embodiment, an injection of over 5,000 chromosomes  
10 reduces or brings the hatch rate to zero. In another embodiment, an injection of over 2,000 chromosomes reduces or brings the hatch rate to zero. For example, an injection of over 1,000 (e.g., 550) chromosomes reduces or brings the hatch rate to zero.

15      For injection, any useful volume of injection buffer may be used for each injection. For example, about 1 nl to about 1  $\mu$ l may be injected. In addition, any useful concentration of chromosomes may be employed in the injection buffer. For example, and without limitation, 1 to about 100,000 chromosomes per microliter may be used. In addition, any useful number of injections may be performed on each egg.

20      In one embodiment, a concentration of 7000-11,500 chromosomes is used per  $\mu$ l of injection buffer (Monteith, D, et al. Methods Mol Biol 240: 227-242, 2004). In one embodiment, 25-100 nanoliters (nl) of injection buffer is used per injection.

25      Any useful avian embryos may be employed in the present invention. For example, the embryos may be collected from 24-36 week-old hens (e.g., commercial White Leghorn variety of *G. gallus*). In one embodiment, a germinal disc is injected with the chromosomes. In one embodiment, the embryo donor hens are inseminated weekly using pooled semen from roosters to produce eggs for injection. Any useful method, such as methods known to those skilled in the art,  
30 may be employed to collect fertilized eggs.

Cytoplasmic injection of artificial chromosomes can be achieved by employing certain microinjection systems or assemblies. In one particularly useful

embodiment, the microinjection assembly or microinjection system disclosed in US Patent Application No. 09/919,143, filed July 31, 2001 (the '143 application), the disclosure of which is incorporated herein in its entirety, is employed. Use of such a cytoplasmic injection device allows for the precise delivery of chromosomes 5 into the cytoplasm of avian embryos, for example, early stage avian embryos, e.g., Stage I embryos.

Typically, following microinjection, the embryos are transferred to the oviduct of recipient hens utilizing any useful technique, such as that disclosed in Olsen, M and Neher, B. J Exp Zool 109: 355-66, 1948, followed by incubation and 10 hatching of the birds.

Any useful method, such as PCR, may be used to test for the production of transchromosomal avians. Typically, the identification of a transchromosomal offspring is confirmed by fluorescence in-situ hybridization (FISH) and/or DNA analysis such as Southern blot or the like. In one useful embodiment, artificial 15 chromosomes can be used as vectors to introduce large DNA payloads, such as nucleotide sequences to be expressed heterologously in the avian to yield a desired biomolecule, of stably maintained genetic information into transgenic chickens. Production of germ-line transchromosomal avians is confirmed by the production of transchromosomal offspring from the G0 birds.

20 The present invention provides for the introduction of desired nucleotide sequences into a chromosome, the chromosome of which can subsequently be isolated/purified and thereafter introduced into an avian as disclosed herein.

A useful chromosome isolation protocol can comprise the steps of inserting 25 a lac-operator sequence (Robinett et al J. Cell Biol. 135: 1685-1700 (1996) into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. In one embodiment, the lac operator region is a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. Insertion can be accomplished, for instance, by identifying a region of known 30 nucleotide sequence associated with a particular avian chromosome. A recombinant DNA molecule may be constructed that comprises the identified region, a recombination site such as attB or attP and a lac-operator concatamer. The recombinant molecule is delivered to an isolated avian cell, for example, but

not limited to, chicken DT40 cells that have elevated homologous recombination activity compared to other avian cell lines, whereupon homologous recombination will integrate the heterologous recombination site and the lac-operator concatamer into the targeted chromosome as shown in the schema illustrated in Fig. 20. A tag-  
5 polypeptide comprising a label domain and a lac repressor domain is also delivered to the cell, for example, by expression from a suitable expression vector. The nucleotide sequence coding for a GFP-lac-repressor fusion protein (Robinett et al,  
J. Cell Biol. 135: 1685-1700 (1996)) may be inserted into the same chromosome as  
the lac-operator insert. The lac repressor sequence, however, can also be within a  
10 different chromosome. An inducible promoter may also be used to allow the  
expression of the GFP-lac-repressor only after chromosome is to be isolated.

Induced expression of the GFP-lac-repressor fusion protein will result in specific binding of the tag fusion polypeptide to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged  
15 mitotic chromosome can be isolated using, for instance, flow cytometry as described in de Jong et al Cytometry 35: 129-133 (1999) and Griffin et al Cytogenet. Cell Genet. 87: 278-281 (1999).

A tagged chromosome can also be isolated using microcell technology requiring treatment of cells with the mitotic inhibitor colcemid to induce the  
20 formation of micronuclei containing intact isolated chromosomes within the cell. Final separation of the micronuclei is then accomplished by centrifugation in cytochalasin as described by Killary & Fournier in Methods Enzymol. 254: 133-152 (1995). Further purification of microcells containing only the desired tagged chromosome could be done by flow cytometry. It is contemplated, however, that  
25 alternative methods to isolate the mitotic chromosomes or microcells, including mechanical isolation or the use of laser scissors and tweezers, and the like.

The present invention envisions the employment of any useful protein-DNA binding or interaction to assist in isolating/purifying chromosomes of the invention. Such other methods in which a desired chromosome can be labeled for  
30 purposes of isolation/purification, are well known in the art including but not limited to, steroid receptor (such as the glucocorticoid receptor):site specific response element systems, see, for example, McNally et al, Science 287:1262-

1265; the bacteriophage lambda repressor system; and human homeobox genes. In addition, certain mutant forms of proteins which are employed in these systems (e.g., mutant proteins which bind their substrate with greater affinity than the non-mutant form of the protein) can be particularly useful for chromosome tagging and 5 subsequent isolation/purification of the chromosomes. Furthermore the invention contemplates the use of a selectable marker to identify cells which contain chromosomes comprising an introduced sequence of interest.

For example, as seen in FIG. 25, an artificial chromosome may include a promoter (e.g., SV40) that will express a marker, such as an antibiotic resistant 10 marker (e.g., hygromycin), when a vector (e.g., plasmid) which includes the gene of interest and the marker coding sequence integrates into the chromosome. For example, a useful cell line such as LMTK- containing the chromosome (A) in FIG. 25 is transfected with the vector B by standard methodologies such as lipofection. After introduction of the vector (B) into the artificial chromosome containing cell 15 line, integration occurs, for example, between integration sites such as lambda attB and attP sites, wherein the hygromycin marker is expressed in the cells which contain the recombined artificial chromosome allowing for selection of the cells. For the employment of such integration sites, integrase or an integrase encoding gene is typically also introduced into the cell. In one useful embodiment, a lambda 20 integrase gene is used which produces an integrase protein with a substitution mutation at the glutamine residue at position 174 to a lysine. This mutation removes the requirement for host factors allowing the integrase to function in cell lines.

This is merely an example of a marker system that can be used to select for 25 chromosomes comprising the nucleotide sequence of interest and other similar systems can be readily envisioned by a practitioner of skill in the art. For example, the method of Gygi et al (2002) Nucleic Acids Res. 30: 2790-2799, the disclosure of which is incorporated by reference herein in its entirety, is contemplated for use in the present invention. Briefly, the protocol provides for the use of synthetic 30 polyamide probes to fluorescently label heterochromatic regions on the chromosomes which are then isolated by flow cytometry. The polyamides bind to the minor groove of DNA of the chromosomes in a sequence specific manner

without the need to disrupt the chromosome (e.g., denature the DNA).

Typically, the artificial chromosomes introduced into avians are stably maintained in the avians and are passed to offspring through the germline. In addition, artificial chromosomes can be stably maintained in avian cell lines such 5 as chicken cell line (DT-40).

The invention is also useful for visualizing gene activity in avian cells as is understood by a practitioner of ordinary skill in the art (See, for example, Tsukamoto, et al (2000) *Nature Cell Biology*, 2:871-878).

Most non-viral methods of gene transfer rely on normal mechanisms used 10 by eukaryotic cells for the uptake and intracellular transport of macromolecules. In certain useful embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject transcriptional regulatory region and operably linked polypeptide-encoding nucleic acid by the targeted cell. Exemplary gene delivery systems of this type include liposomal 15 derived systems, poly-lysine conjugates, and artificial viral envelopes. Modified chromosomes as described above may be delivered to isolated avian embryonic cells for subsequent introduction to an embryo.

In a representative embodiment, a nucleic acid molecule can be entrapped 20 in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al, 1992, NO Shinkei Geka 20: 547-551; PCT publication WO91/06309, published May 16, 1991; Japanese patent application 1047381, published February 21, 1989; and European patent publication EP-A-43075, published January 6, 1982, all of which are incorporated herein by reference in 25 their entireties).

In similar fashion, the gene delivery system can comprise an antibody or cell surface ligand that is cross-linked with a gene binding agent such as polylysine (see, for example, PCT publications WO93/04701, published March 18, 1993; WO92/22635, published December 23, 1992; WO92/20316, published November 30 26, 1992; WO92/19749, published November 12, 1992; and WO92/06180, published April 16, 1992, the disclosures of which are incorporated herein by reference in their entireties). It will also be appreciated that effective delivery of

the subject nucleic acid constructs via receptor-mediated endocytosis can be improved using agents which enhance escape of genes from the endosomal structures. For instance, whole adenovirus or fusogenic peptides of the influenza HA gene product can be used as part of the delivery system to induce efficient  
5 disruption of DNA-containing endosomes (Mulligan et al, 1993, Science 260:926-932; Wagner et al, 1992, Proc. Natl. Acad. Sci. 89:7934-7938; and Christiano et al, 1993, Proc. Natl. Acad. Sci. 90:2122-2126, all of which are incorporated herein by reference in their entireties). It is further contemplated that a recombinant nucleic acid molecule of the present invention may be delivered to a target host cell by  
10 other non-viral methods including by gene gun, microinjection, sperm-mediated transfer, or the like.

In one embodiment of the invention, an expression vector that comprises a recombination site, such as an attB site, and a region encoding a polypeptide deposited into an egg white are delivered to oviduct cells by in vivo electroporation. In this method, the luminal surface of an avian oviduct is surgically exposed. A buffered solution of the expression vector and a source of integrase activity such as a second expression vector expressing integrase (for example, pCMV-int) is deposited on the luminal surface. Electroporation electrodes are then positioned on either side of the oviduct wall, the luminal  
15 electrode contacting the expression vector solution. After electroporation, the surgical incisions are closed. The electroporation will deliver the expression vectors to some, if not all, treated recipient oviduct cells to create a tissue-specific chimeric animal. Expression of the integrase allows for the integration of the heterologous polynucleotide into the genome of recipient oviduct cells. While this  
20 method may be used with any bird, a useful recipient is a chicken due to the size of the oviduct. Also useful is a transgenic bird that has a transgenic attP recombinant site in the nuclear genomes of recipient oviduct cells, thus increasing the efficiency  
25 of integration of the expression vector.

The attB/P integrase system is useful in the in vivo electroporation method  
30 to allow the formation of stable genetically transformed oviduct cells that otherwise progressively lose the heterologous expression vector.

The stably modified oviduct cells will express the heterologous

polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

5 Another aspect of the invention is the generation of a trisomic or transchromosomal avian cell comprising a genetically modified extra chromosome. The extra chromosome may be an artificial chromosome or an isolated avian chromosome that has been genetically modified. Introduction of the extra chromosome to an avian cell will generate a trisomic or transchromosomal cell  
10 with  $2n+1$  chromosomes, where n is the haploid number of chromosomes of a normal avian cell.

Delivery of an isolated chromosome into an isolated avian cell or embryo can be accomplished in several ways. Isolated mitotic chromosomes or a micronucleus containing an interphase chromosome can be injected into early stage  
15 I embryos by cytoplasmic injection. The injected zygote would then be surgically transferred to a recipient hen for the production and laying of a hard shell egg. This hard shell egg would then be incubated until hatching of a chick.

In one embodiment, isolated microcells which contain the artificial chromosome can be fused to primordial germ cells (PGCs) isolated from the blood stream of late stage 15 embryos as described by Killary & Fournier in Methods Enzymol. 254: 133-152 (1995). The PGC/microcell hybrids can then be transplanted into the blood stream of a recipient embryo to produce germline chimeric chickens. (See Naito et al, Mol. Reprod. Dev. 39: 153-161 (1994)). The manipulated eggs would then be incubated until hatching of the bird.

25 Blastodermal cells isolated from stage X embryos can be transfected with isolated mitotic chromosomes. Following in vitro transfection, the cells are transplanted back into stage X embryos as described, for example, in Etches et al, Poult. Sci., 72: 882-889 (1993), and the manipulated eggs are incubated to hatching.

30 Stage X blastodermal cells can also be fused with isolated microcells and then transplanted back into stage X embryos or fused to somatic cells to be used as nuclear donors for nuclear transfer as described by Kuroiwa et al, Nat.

Biotechnol. 20: 889-894 (2002).

Chromosomal vectors, as described above, may be delivered to a recipient avian cell by, for example, microinjection, liposomal delivery or microcell fusion.

In the methods of the invention, a site-specific integrase is introduced into 5 an avian cell whose genome is to be modified. Methods of introducing functional proteins into cells are well known in the art. Introduction of purified integrase protein can ensure a transient presence of the protein and its activity. Thus, the lack of permanence associated with most expression vectors is not expected to be detrimental.

10 The integrase used in the practice of the present invention can be introduced into a target cell before, concurrently with, or after the introduction of a site-specific vector. The integrase can be directly introduced into a cell as a protein, for example, by using liposomes, coated particles, or microinjection, or into the blastodermal layer of an early stage avian embryo by microinjection. A source of 15 the integrase can also be delivered to an avian cell by introducing to the cell an mRNA encoding the integrase and which can be expressed in the recipient cell as an integrase polypeptide. Alternately, a DNA molecule encoding the integrase can be introduced into the cell using a suitable expression vector.

The present invention provides novel nucleic acid vectors and methods of 20 use that allow integrases, such as phiC31 integrase, to efficiently integrate a heterologous nucleic acid into a vertebrate animal genome, for example, an avian genome. A novel finding is that the phiC31 integrase is remarkably efficient in 25 avian cells and increases the rate of integration of heterologous nucleic acid at least 30-fold over that of random integration. Furthermore, the phiC31 integrase works equally well at 37°C and 41°C, indicating that it will function in the environment of the developing avian embryo, as shown in Example 1.

It is important to note that the present invention is not bound by any mechanism or theory of operation. For example, the mechanism by which 30 integrase, or any other substance described herein, facilitates transgenesis is unimportant. Integrase, for example, may facilitate transgenesis by mediating the integration of DNA into the genome of a recipient cell or integrase may facilitate transgenesis by facilitating the entry of the DNA into the cell or integrase may

facilitate transgenesis by some other mechanism.

The site-specific vector components described above are useful in the construction of expression cassettes containing sequences encoding an integrase. One integrase-expressing vector useful in the methods of the invention is pCMV-C31int (SEQ ID NO: 1 as shown in Fig. 9) where the phiC31 integrase is encoded by a region under the expression control of the strong CMV promoter. Another useful promoter is the RSV promoter as used in SEQ ID NO: 9 shown in Fig. 17. Expression of the integrase is typically desired to be transient. Accordingly, vectors providing transient expression of the integrase are useful. However, expression of the integrase can be regulated in other ways, for example, by placing the expression of the integrase under the control of a regulatable promoter (i.e., a promoter whose expression can be selectively induced or repressed).

Delivery of the nucleic acids introduced into cells, for example, embryonic cells (e.g., avian cells), using methods of the invention may also be enhanced by mixing the nucleic acid to be introduced with a nuclear localization signal (NLS) peptide prior to introduction, for example, microinjection, of the nucleic acid. Nuclear localization signal (NLS) sequences are a class of short amino acid sequences which may be exploited for cellular import of linked cargo into a nucleus. The present invention envisions the use of any useful NLS peptide, including but not limited to, the NLS peptide of SV40 virus T-antigen.

An NLS of the invention is an amino acid sequence which mediates nuclear transport into the nucleus, wherein deletion of the NLS reduces transport into the nucleus. In certain embodiments, an NLS is a cationic peptide, for example, a highly cationic peptide. The present invention includes the use of any NLS sequence, including but not limited to, SV40 virus T-antigen. NLSs known in the art include, but are not limited to those discussed in Cokol et al, 2000, EMBO Reports, 1(5):411-415, Boulikas, T., 1993, Crit. Rev. Eukaryot. Gene Expr., 3:193-227, Collas, P. et al, 1996, Transgenic Research, 5: 451-458, Collas and Alestrom, 1997, Biochem. Cell Biol. 75: 633-640, Collas and Alestrom, 1998, Transgenic Research, 7: 303-309, Collas and Alestrom, Mol. Reprod. Devel., 1996, 45:431-438. The disclosure of each of these references is incorporated by reference herein in its entirety.

Not to be bound by any mechanism of operation, DNA is protected and hence stabilized by cationic polymers. The stability of DNA molecules in the cytoplasm of cells may be increased by mixing the DNA to be introduced, for example, microinjected with cationic polymers (for example, branched cationic polymers), such as polyethylenimine (PEI), polylysine, DEAE-dextran, starburst dendrimers, starburst polyamidoamine dendrimers, and other materials that package and condense the DNA molecules (Kukowska-Latallo et al, 1996, Proc. Natl. Acad. Sci. USA 93:4897-4902).

Once the DNA molecules are delivered to the cytoplasm of cells, they migrate into the cell's endocytotic vesicles. Furthermore, migration into the cell's endosome is followed by fast inactivation of DNA within the endolysosomal compartment in transfected or injected cells, both *in vitro* and *in vivo* (Godbey, W, et al 1999, Proc Natl Acad Sci U S A 96: 5177-5181; and Lechardeur, D, et al 1999, Gene Ther 6: 482-497; and references cited therein). Accordingly, in certain embodiments, DNA uptake is enhanced by the receptor-mediated endocytosis pathway using transferrin-polylysine conjugates or adenoviral-mediated vesicle disruption to effect the release of DNA from endosomes. However, the invention is not limited to this or any other theory or mechanism of operation referred to herein.

Buffering the endosomal pH using endosomal-scaping elements also protects DNA from degradation (Kircheis, R, et al 2001, Adv Drug Deliv Rev 53: 341-358 ; Boussif, O, et al 1995, Proc Natl Acad Sci U S A 92: 7297-7301; and Pollard, H, et al 1998, J Biol Chem 273: 7507-7511; and references cited therein). Thus, in certain embodiments, DNA complexes are delivered with polycations or cationic polymers that possess substantial buffering capacity below physiological pH, such as polyethylenimine, lipopolyamines and polyamidoamine polymers. In certain embodiments, DNA condensing compounds, such as the ones described above, are combined with viruses (Curiel, D, et al Proc Natl Acad Sci U S A 88: 8850-8854, 1991; Wagner, E, et al Proc Natl Acad Sci U S A 89: 6099-6103, 1992 and Cotten, M, et al, 1992, Proc Natl Acad Sci U S A 89: 6094-6098), viral peptides (Wagner, E, et al 1992, Proc Natl Acad Sci U S A 89: 7934-7938; Plank, C, et al 1994, J Biol Chem 269: 12918-12924) and subunits of toxins (Uhrek, C,

et al, 1998, J Biol Chem 273: 8835-48). These materials significantly enhance the release of DNA from endosomes. In certain embodiments, viruses, viral peptides, toxins or subunits of toxins may be coupled to DNA/polylysine complexes via biochemical means or specifically by a streptavidin-biotin bridge (Wagner et al, 5 1992, Proc. Natl. Acad. Sci. USA 89:6099-6103; Plank et al, 1994, J. Biol Chem. 269(17):12918-12924). In other certain embodiments, the virus that is complexed with the DNA may be adenovirus, retrovirus, vaccinia virus, or parvovirus. The viruses may be linked to PEI or another cationic polymer associated with the nucleic acid. In certain embodiments, the virus may be alphavirus, 10 orthomyxovirus, or picornavirus. In certain embodiments, the virus is defective or chemically inactivated. The virus may be inactivated by short-wave UV radiation or the DNA intercalator psoralen plus long-wave UV. The adenovirus may be coupled to polylysine, either enzymatically through the action of transglutaminase or biochemically by biotinylation adenovirus and streptavidinylating the polylysine 15 moiety. Transferrin may also be useful in combination with cationic polymers, adenoviruses and/or other materials disclosed herein to produce transgenic avians. For example, DNA complexes containing PEI, PEI-modified transferrin, and PEI-bound influenza peptides may be used to enhance transgenic avian production.

In other certain embodiments, complexes containing plasmid DNA, 20 transferrin-PEI conjugates, and PEI-conjugated peptides derived from the N-terminal sequence of the influenza virus hemagglutinin subunit HA-2 may be used to produce transgenic chickens. In certain embodiments, the PEI-conjugated peptide may be an amino-terminal amino acid sequence of influenza virus hemagglutinin which may be elongated by an amphipathic helix or by carboxyl-terminal dimerization. 25

The present invention provides for methods of dispersing or distributing nucleic acid in a cell, for example, in an avian cell. The avian cell may be, for example, and without limitation, a cell of a stage I avian embryo, a cell of a stage II avian embryo, a cell of a stage III avian embryo, a cell of a stage IV avian embryo, 30 a cell of a stage V avian embryo, a cell of a stage VI avian embryo, a cell of a stage VII avian embryo, a cell of a stage VIII avian embryo, a cell of a stage IX avian embryo, a cell of a stage X avian embryo, a cell of a stage XI avian embryo or a

cell of a stage XII avian embryo. In one particularly useful embodiment, the avian cell is a cell of a stage X avian embryo.

In one aspect of the present invention, cationic polymers are useful to distribute, for example, homogeneously distribute, nucleic acid introduced into a 5 cell, for example, an embryonic avian cell. The present invention contemplates the use of cationic polymers including, but not limited to, those disclosed herein.

However, substances other than cationic polymers also capable of distributing or dispersing nucleic acids in a cell are included within the scope of the present invention.

10       The concentration of cationic polymer used is not critical though, in one useful embodiment, enough cationic polymer is present to coat the nucleic acid to be introduced into the avian cell. The cationic polymer may be present in an aqueous mixture with the nucleic acid to be introduced into the cell at a concentration in a range of an amount equal to about the weight of the nucleic acid  
15       to a concentration wherein the solution is saturated with cationic polymer. In one useful embodiment, the cationic polymer is present in an amount in a range of about 0.01% to about 50 %, for example, about 0.1% to about 20% (e.g., about 5%). The molecular weights of the cationic polymers can range from a molecular weight of about 1,000 to a molecular weight of about 1,000,000. In one  
20       embodiment, the molecular weight of the cationic polymers range from about 5,000 to about 100,000 for example, about 20,000 to about 30,000.

In one particularly useful aspect of the invention, procedures that are effective to facilitate the production of a transgenic avian may be combined to provide for an enhanced production of a transgenic avian wherein the enhanced  
25       production is an improved production of a transgenic avian relative to the production of a transgenic avian by only one of the procedures employed in the combination. For example, one or more of integrase activity, NLS, cationic polymer or other technique useful to enhance transgenic avian production disclosed herein can be used in the same procedure to provide for an enhanced production of  
30       transgenic avians relative to an identical procedure which does not employ all of the same techniques useful to enhance transgenic avian production.

Another aspect of the present invention is a vertebrate animal cell which has been genetically modified with a transgene vector according to the present invention and as described herein. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically 5 transformed cell line, including a sustainable cell line. The transfected cell according to the present invention may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration of a transgene comprises a 10 heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

A retroviral vector can be used to deliver a recombination site such as an att site into the cellular genomes, such as avian genomes, since an attP or attB site is 15 less than 300 bp. For example, the attP site can be inserted into the NLB retroviral vector, which is based on the avian leukosis virus genome. A lentiviral vector is a particularly suitable vector because lentiviral vectors can transduce non-dividing cells, so that a higher percentage of cells will have an integrated attP site.

The lacZ region of NLB is replaced by the attP sequence. A producer cell 20 line would be created by transformation of, for example, the Isolde cell line capable of producing a packaged recombinant NLB-attP virus pseudo-typed with the envA envelope protein. Supernatant from the Isolde NLB-attP line is concentrated by centrifugation to produce high titer preparations of the retroviral vector that can then be used to deliver the attP site to the genome of a cell, for example, as 25 described in Example 9 below.

In one embodiment, an attP-containing line of transgenic birds are a source of attP transgenic embryos and embryonic cells. Fertile zygotes and oocytes bearing a heterologous attP site in either the maternal, paternal, or both, genomes can be used for transgenic insertion of a desired heterologous polynucleotide. A 30 transgene vector bearing an attB site, for example, would be injected into the cytoplasm along with either an integrase expression plasmid, mRNA encoding the integrase or the purified integrase protein. The oocyte or zygote is then cultured to

hatch by ex ovo methods or reintroduced into a recipient hen such that the hen lays a hard shell egg the next day containing the injected egg.

In another example, fertile stage I to XII embryos, for example, stage VII to XII embryos, hemizygous or homozygous for the heterologous integration site, for 5 example, the attP sequence, may be used as a source of blastodermal cells. The cells are harvested and then transfected with a transgene vector bearing a second recombination site, such as an attB site, plus a nucleotide sequence of interest along with a source of integrase. The transfected cells are then injected into the 10 subgerminal cavity of windowed fertile eggs. The chicks that hatch will bear the nucleotide sequence of interest and the second integration site integrated into the attP site in a percentage of their somatic and germ cells. To obtain fully transgenic birds, chicks are raised to sexual maturity and those that are positive for the transgene in their semen are bred to non-transgenic mates. As disclosed herein, in certain embodiments, the cells of the invention, e.g., embryos, may include an 15 integrase which specifically recognizes recombination sites and which is introduced into cells containing a nucleic acid construct of the invention under conditions such that the nucleic acid sequence(s) of interest will be inserted into the nuclear genome. Methods for introducing such an integrase into a cell are described herein. In some embodiments, the site-specific integrase is introduced 20 into the cell as a polypeptide. In alternative embodiments, the site-specific integrase is introduced into the transgenic cell as a polynucleotide encoding the integrase, such as an expression cassette optionally carried on a transient expression vector, and comprising a polynucleotide encoding the recombinase.

In one embodiment, the invention is directed to methods of using a vector 25 for site-specific integration of a heterologous nucleotide sequence into the genome of a cell, the vector comprising a circular backbone vector, a polynucleotide of interest operably linked to a promoter, and a first recombination site, wherein the genome of the cell comprises a second recombination site and recombination between the first and second recombination sites is facilitated by an integrase. In 30 certain embodiments, the integrase facilitates recombination between a bacterial genomic recombination site (attB) and a phage genomic recombination site (attP).

In another embodiment, the invention is directed to a cell having a transformed genome comprising an integrated heterologous polynucleotide of interest whose integration, mediated by an integrase, was into a recombination site native to the cell genome and the integration created a recombination-product site 5 comprising the polynucleotide sequence. In yet another embodiment, integration of the polynucleotide was into a recombination site not native to the cell genome, but instead into a heterologous recombination site engineered into the cell genome.

In further embodiments, the invention is directed to transgenic vertebrate animals, such as transgenic birds, comprising a modified cell and progeny thereof 10 as described above, as well as methods of producing the same.

For example, cells genetically modified to carry a heterologous attB or attP site by the methods of the present invention can be maintained under conditions that, for example, keep them alive but do not promote growth and/or cause the cells to differentiate or dedifferentiate. Cell culture conditions may be permissive for 15 the action of the integrase in the cells, although regulation of the activity of the integrase may also be modulated by culture conditions (e.g., raising or lowering the temperature at which the cells are cultured).

One aspect of the invention are methods for generating a genetically modified cell for example, an avian cell, and progeny thereof, using a tagged 20 chromosome. The methods may include providing an isolated modified chromosome comprising a lac operator region and a first recombination site, delivering the modified chromosome to a avian cell, thereby generating a trisomic or transchromosomal avian cell, delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain, 25 delivering a source of integrase activity to the avian cell, delivering a polynucleotide comprising a second recombination site and a region encoding a polypeptide to the avian cell, maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified 30 chromosome and generating a genetically modified avian cell, expressing the tag polypeptide by the avian cell, allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome, and isolating the modified

chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

In one embodiment of the invention, the second avian cell is selected from the group consisting of a stage VII-XII blastodermal cell, a stage I embryo, a stage 5 X embryo; an isolated primordial germ cell, an isolated non-embryonic cell, and an oviduct cell.

In various embodiments, the isolated modified chromosome is an avian chromosome or an artificial chromosome.

In other embodiments of the invention, the step of providing an isolated 10 modified chromosome comprising a lac operator region and a first recombination site comprises the steps of generating a trisomic or transchromosomal avian cell by delivering to an isolated avian cell an isolated chromosome and a polynucleotide comprising a lac operator and a second recombination site, maintaining the trisomic or transchromosomal cell under conditions whereby the heterologous 15 polynucleotide is integrated into the chromosome by homologous recombination, delivering to the avian cell a source of a tag polypeptide to label the chromosome, and isolating the labeled chromosome.

In one embodiment of the invention, the lac operator region is a concatamer 20 of lac operators. In other embodiments of the invention, the tag polypeptide is expressed from an expression vector.

In one embodiment of the invention, the tag polypeptide is microinjected into the cell. In various embodiments of the invention, the method of delivery of a 25 chromosome to an avian cell is selected from the group consisting of liposome delivery, microinjection, microcell, electroporation and gene gun delivery, or a combination thereof.

In embodiments of the invention, the fluorescent domain of the tag polypeptide is GFP.

In one embodiment of the invention, the method further comprises the step 30 of delivering the second avian cell to an avian embryo. The embryo may be maintained under conditions suitable for hatching as a chick.

In one embodiment of the invention, the second avian cell is maintained under conditions suitable for the proliferation of the cell, and progeny thereof.

In various embodiments of the invention, the source of integrase activity is delivered to a first avian cell as a polypeptide or expressed from a polynucleotide, said polynucleotide being selected from an mRNA and an expression vector.

In one embodiment of the invention, the tag polypeptide activity is  
5 delivered to the avian cell as a polypeptide or expressed from a polynucleotide operably linked to a promoter. In another embodiment of the invention, the promoter is an inducible promoter. In yet another embodiment of the invention, the integrase is phiC31 integrase and in various embodiments of the invention, the first and second recombination sites are selected from an attB and an attP site, but  
10 wherein the first and second sites are not identical.

Other aspects of the present invention include methods of expressing a heterologous polypeptide in vertebrate cells by stably transfecting cells using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described herein, and culturing the transfected cells under conditions suitable for expression  
15 of the heterologous polypeptide. In addition, the present invention includes methods of expressing a heterologous polypeptide in a transgenic vertebrate animal by producing a transgenic vertebrate animal using methods known in the field or described herein in combination with using site-specific integration of nucleic acid molecules as described herein, and exposing the animal to conditions suitable for  
20 expression of the heterologous polypeptide.

The protein of the present invention may be produced in purified form by any known conventional techniques. For example, in the case of heterologous protein production in eggs, the egg white may be homogenized and centrifuged. The supernatant may then be subjected to sequential ammonium sulfate  
25 precipitation and heat treatment. The fraction containing the protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by HPLC or other methods well known in the art of protein purification.

30 The methods of the invention are useful for expressing nucleic acid sequences that are optimized for expression in the host cells and which encode desired polypeptides or derivatives and fragments thereof. Derivatives include, for

instance, polypeptides with conservative amino acid replacements, that is, those within a family of amino acids that are related in their side chains (commonly known as acidic, basic, nonpolar, and uncharged polar amino acids). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids and other groupings are known in the art (see, for example, "Biochemistry", 2nd ed, L. Stryer, ed., W.H. Freeman & Co., 1981). Peptides in which more than one replacement has taken place can readily be tested for activity in the same manner as derivatives with a single replacement, using conventional polypeptide activity assays (e.g. for enzymatic or ligand binding activities).

Regarding codon optimization, if the recombinant nucleic acid molecules are transfected into a recipient chicken cell, the sequence of the nucleic acid insert to be expressed can be optimized for chicken codon usage. This may be determined from the codon usage of at least one, or more than one, protein expressed in a chicken cell according to well known principles. For example, in the chicken the codon usage could be determined from the nucleic acid sequences encoding the proteins such as lysozyme, ovalbumin, ovomucin and ovotransferrin of chicken. Optimization of the sequence for codon usage can elevate the level of translation in avian eggs.

The present invention provides methods for the production of a protein by cells comprising the steps of maintaining a cell, transfecting with a first expression vector and, optionally, a second expression vector, under conditions suitable for proliferation and/or gene expression and such that an integrase will mediate site specific recombination at att sites. The expression vectors may each have a transcription unit comprising a nucleotide sequence encoding a heterologous polypeptide, wherein one polypeptide is an integrase, a transcription promoter, and a transcriptional terminator. The cells may then be maintained under conditions for the expression and production of the desired heterologous polypeptide(s).

The present invention further relates to methods for gene expression by cells, such as avian cells, from nucleic acid vectors, and transgenes derived therefrom, that include more than one polypeptide-encoding region wherein, for example, a first polypeptide-encoding region can be operatively linked to an avian promoter and a second polypeptide-encoding region is operatively linked to an

Internal Ribosome Entry Sequence (IRES). It is contemplated that the first polypeptide-encoding region, the IRES and the second polypeptide-encoding region of a recombinant DNA of the present invention may be arranged linearly, with the IRES operably positioned immediately 5' of the second polypeptide-  
5 encoding region. This nucleic acid construct can be used for the production of certain proteins in vertebrate animals or in their cells. For example, when inserted into the genome of an avian cell or a bird and expressed therein, will generate individual polypeptides that may be post-translationally modified and combined in the white of a hard shell bird egg. Alternatively, the expressed polypeptides may  
10 be isolated from an avian egg and combined in vitro.

The invention, therefore, includes methods for producing multimeric proteins including immunoglobulins, such as antibodies, and antigen binding fragments thereof. Thus, in one embodiment of the present invention, the multimeric protein is an immunoglobulin, wherein the first and second  
15 heterologous polypeptides are immunoglobulin heavy and light chains respectively. Illustrative examples of this and other aspects of the present invention for the production of heterologous multimeric polypeptides in avian cells are fully disclosed in U.S. Patent Application No. 09/877,374, filed June 8, 2001, and U.S. Patent Application No. 10/251,364, filed September 18, 2002, both of which are  
20 incorporated herein by reference in their entirety.

Accordingly, the invention further provides immunoglobulin and other multimeric proteins that have been produced by transgenic vertebrates including avians of the invention.

In various embodiments, an immunoglobulin polypeptide encoded by the  
25 transcriptional unit of at least one expression vector may be an immunoglobulin heavy chain polypeptide comprising a variable region or a variant thereof, and may further comprise a D region, a J region, a C region, or a combination thereof. An immunoglobulin polypeptide encoded by an expression vector may also be an immunoglobulin light chain polypeptide comprising a variable region or a variant  
30 thereof, and may further comprise a J region and a C region. The present invention also contemplates multiple immunoglobulin regions that are derived from the same animal species, or a mixture of species including, but not only, human, mouse, rat,

rabbit and chicken. In certain embodiments, the antibodies are human or humanized.

In other embodiments, the immunoglobulin polypeptide encoded by at least one expression vector comprises an immunoglobulin heavy chain variable region, 5 an immunoglobulin light chain variable region, and a linker peptide thereby forming a single-chain antibody capable of selectively binding an antigen.

Examples of therapeutic antibodies that may be produced in methods of the invention include but are not limited to HERCEPTIN™ (Trastuzumab) (Genentech, CA) which is a humanized anti-HER2 monoclonal antibody for the 10 treatment of patients with metastatic breast cancer; REOPRO™ (abciximab) (Centocor) which is an anti-glycoprotein IIb/IIIa receptor on the platelets for the prevention of clot formation; ZENAPAX™ (daclizumab) (Roche Pharmaceuticals, Switzerland) which is an immunosuppressive, humanized anti-CD25 monoclonal antibody for the prevention of acute renal allograft rejection; PANOREX™ which 15 is a murine anti-l7-IA cell surface antigen IgG2a antibody (Glaxo Wellcome/Centocor); BEC2 which is a murine anti-idiotype (GD3 epitope) IgG antibody (ImClone System); IMC-C225 which is a chimeric anti-EGFR IgG antibody (ImClone System); VITAXIN™ which is a humanized anti- $\alpha$ V $\beta$ 3 integrin antibody (Applied Molecular Evolution/MedImmune); Campath 1H/LDP-03 which 20 is a humanized anti CD52 IgG1 antibody (Leukosite); Smart M195 which is a humanized anti-CD33 IgG antibody (Protein Design Lab/Kanebo); RITUXANTM which is a chimeric anti-CD20 IgG1 antibody (IDEC Pharm/Genentech, Roche/Zettyaku); LYMPHOCIDE™ which is a humanized anti-CD22 IgG antibody (Immunomedics); ICM3 is a humanized anti-ICAM3 antibody (ICOS 25 Pharm); IDEC-114 is a primate anti-CD80 antibody (IDEC Pharm/Mitsubishi); ZEVALINTM is a radiolabelled murine anti-CD20 antibody (IDEC/Schering AG); IDEC-131 is a humanized anti-CD40L antibody (IDEC/Eisai); IDEC-151 is a primatized anti-CD4 antibody (IDEC); IDEC-152 is a primatized anti-CD23 antibody (IDEC/Seikagaku); SMART anti-CD3 is a humanized anti-CD3 IgG 30 (Protein Design Lab); 5G1.1 is a humanized anti-complement factor 5 (CS) antibody (Alexion Pharm); D2E7 is a humanized anti-TNF- $\alpha$  antibody (CATIBASF); CDP870 is a humanized anti-TNF- $\alpha$  Fab fragment (Celltech);

IDEc-151 is a primatized anti-CD4 IgG1 antibody (IDEc Pharm/SmithKline Beecham); MDX-CD4 is a human anti-CD4 IgG antibody (Medarex/Eisai/Genmab); CDP571 is a humanized anti-TNF- $\alpha$  IgG4 antibody (Celltech); LDP-02 is a humanized anti- $\alpha$ 4 $\beta$ 7 antibody (LeukoSite/Genentech); OrthoClone OKT4A is a humanized anti-CD4 IgG antibody (Ortho Biotech); ANTOVA™ is a humanized anti-CD40L IgG antibody (Biogen); ANTEGRENTM is a humanized anti-VLA-4 IgG antibody (Elan); and CAT-152 is a human anti-TGF- $\beta$ 2 antibody (Cambridge Ab Tech).

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and  $\beta$ -casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention. Other possible proteins to be produced include, but are not limited to, albumin,  $\alpha$ -1 antitrypsin, antithrombin III, collagen, factors VIII, IX, X (and the like), fibrinogen, hyaluronic acid, insulin, lactoferrin, protein C, erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), tissue-type plasminogen activator (tPA), feed additive enzymes, somatotropin, and chymotrypsin. Immunoglobulins (shown, for example in Example 10 below) and genetically engineered antibodies, including immunotoxins which bind to surface antigens on human tumor cells and destroy them, can also be expressed for use as pharmaceuticals or diagnostics.

Other specific examples of therapeutic substances which are contemplated for production as disclosed herein include, with out limitation, Factor VIII (e.g., Recombinate®, Bioclate®, Kogenate®, Helixate® (Centeon), B-domain deleted Factor VIII (e.g., ReFacto®), Factor VIIa (e.g., NovoSeven®), Factor IX (e.g., Benefix®), anticoagulant; recombinant hirudin (e.g., Revasc®, Refludan®) Alteplase, tPA (e.g., Activase®), Reteplase, tPA, tPA – 3 of 5 domains deleted, Ecokinase®, Retavase®, Rapilysin®, insulin (e.g., Humulin®, Novolin®, Insuman®) insulin lispro (e.g., Humalog®), Bio Lysprol, Liprolog®, insulin Aspart, iNovoRapid®, insulin glargine, long-acting insulin analog (e.g., Lantus®),

rhGH (e.g., Protropin®, Humatropin®, Nutropin®, BioTropin®, Genotropin®, Norditropin®, Saizen®, Serostim®), glucagons (e.g., Glucagen®), TSH (e.g., Thyrogen®, Gonal F®, Puregon®), follitropin-beta FSH (e.g., Folliestim®), EPO (e.g., Epogen®, Procrit®, Neorecormon®), GM-CSF (e.g., Leukine®,  
5 Neupogen®), PDGH (e.g., Regranex®), IFN alpa2a (e.g., Roferon A®), INF-apha (e.g., Infergen®), IFN alpa2b (e.g., Intron A®, Alfatronol®, Virtron®), ribavirin & INF-alpha 2b (e.g., Robetron®) INF-beta 1b, differs from h protein by C17 to S (e.g., Betaferon®), IFN-beta 1a (e.g., Avonex®, Rebif®), IFN-gamma1b (e.g., Actimmune®), IL-2 (e.g., Proleukin®) rIL-11 (e.g., Neumega®), rHBsAg (e.g.,  
10 Recombivax®), Combination vaccine containing HBsAgn as one component (e.g., Comvax®, Tritarix®, Twinrix®, Primavax®, Procomax®), OspA, a lipoprotein found on the surface of *B. burgdorferi* (e.g., Lymerix®), murine MAb directed against t-lymphocyte antigen CD3 (e.g., Orthoclone OKT3®), murine MAb directed against TAG-72, tumor-associated glycoprotein (e.g., OncoScint CR/OV®), FAB  
15 fragments derived from chimeric MAb, directed against platelet surface receptor GPII(b)/III(a) (e.g., ReoPro®), murine MAb fragment directed against tumor-associated antigen CA125 (e.g., Indimacis®), murine MAb fragment directed against human carcinoembryonic antigen, CEA (e.g., CEA-scan®), murine MAb fragment directed against human cardiac myosin (e.g., MyoScint®), murine MAb  
20 fragment directed against tumor surface antigen PSMA (e.g., ProstaScint®), murine MAb fragments (FAB/FAb2 mix) directed against HMW-MAA (e.g., Tacnemab®), murine MAb fragment (FAB) directed against carcinoma-associated antigen (e.g., Verluma®), MAb fragments (FAB) directed against NCA 90, a surface granulocyte nonspecific cross reacting antigen (e.g., LeukoScan®),  
25 chimeric MAb directed against CD20 antigen found on surface of B lymphocytes (e.g., Rituxan®), humanized MAb directed against the alpha chain of the IL2 receptor (e.g., Zenapax®), chimeric MAb directed against the alpha chain of the IL2 receptor (e.g., Simulect®), chimeric MAb directed against TNF-alpha (e.g., Remicade®), humanized MAb directed against an epitope on the surface of  
30 respiratory syncytial virus (e.g., Synagis®), humanized MAb directed against HER 2, i.e., human epidermal growth factor receptor 2 (e.g., Herceptin®), human MAb directed against cytokeratin tumor-associated antigen (e.g., Humaspect®), anti-

CTLA4, chimeric MAb directed against CD 20 surface antigen of B lymphocytes (e.g., Mabthera®), dornase-alpha DNase (e.g., Pulmozyme®), beta glucocerebrosidase (e.g., Cerezyme®), TNF-alpha (e.g., Beromun®), IL-2-diphtheria toxin fusion protein that targets cells displaying a surface IL-2 receptor  
5 (e.g., Ontak®), TNFR-IgG fragment fusion protein (e.g., Enbrel®), Laronidase, Recombinant DNA enzyme, (e.g., Aldurazyme®), Alefacept, Amevive®, Darbepoetin alfa (Colony stimulating factor) (e.g., Aranesp®), Tositumomab and iodine 131 tositumomab, murine MAb, Bexxar®, Alemtuzumab, Campath®, Rasburicase, Elitek®), Agalsidase beta, Fabrazyme®, FluMist®, Teriparatide,  
10 Parathyroid hormone derivative (e.g., Forteo®), Enfuvirtide Fuzeon®, Adalimumab (IgG1) (e.g., Humira®), Anakinra, Biological modifier (e.g., Kineret®), nesiritide, Human B-type natriuretic peptide (hBNP) (e.g., Natrecor®), Pegfilgrastim, Colony stimulating factor (e.g., Neulasta®), ribavarin and peg Intron A (e.g., Rebetron®), Pegvisomant, PEGylated human growth hormone  
15 receptor antagonist, (e.g., Somavert®), recombinant activated protein C (e.g., Xigris®), Omalizumab, Immunoglobulin E (IgE) blocker (e.g., Xolair®) and Ibritumomab tiuxetan (murine MAb) (e.g., Zevalin®).

In various embodiments of the transgenic vertebrate animal of the present invention, the expression of the transgene may be restricted to specific subsets of  
20 cells, tissues or developmental stages utilizing, for example, trans-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene  
25 in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Another aspect of the present invention provides a method for the production of a heterologous protein capable of forming an antibody suitable for  
30 selectively binding an antigen. This method comprises a step of producing a transgenic vertebrate animal incorporating at least one transgene, the transgene encoding at least one heterologous polypeptide selected from an immunoglobulin

heavy chain variable region, an immunoglobulin heavy chain comprising a variable region and a constant region, an immunoglobulin light chain variable region, an immunoglobulin light chain comprising a variable region and a constant region, and a single-chain antibody comprising two peptide-linked immunoglobulin 5 variable regions.

In one embodiment of this method, the isolated heterologous protein is an antibody capable of selectively binding to an antigen and which may be generated by combining at least one immunoglobulin heavy chain variable region and at least one immunoglobulin light chain variable region, for example, cross-linked by at 10 least one disulfide bridge. The combination of the two variable regions generates a binding site that binds an antigen using methods for antibody reconstitution that are well known in the art.

The present invention also encompasses immunoglobulin heavy and light chains, or variants or derivatives thereof, to be expressed in separate transgenic 15 avians, and thereafter isolated from separate media including serum or eggs, each isolate comprising one or more distinct species of immunoglobulin polypeptide. The method may further comprise the step of combining a plurality of isolated heterologous immunoglobulin polypeptides, thereby producing an antibody capable of selectively binding to an antigen. In this embodiment, for instance, two or more 20 individual transgenic avians may be generated wherein one transgenic produces serum or eggs having an immunoglobulin heavy chain variable region, or a polypeptide comprising such, expressed therein. A second transgenic animal, having a second transgene, produces serum or eggs having an immunoglobulin light chain variable region, or a polypeptide comprising such, expressed therein. 25 The polypeptides from two or more transgenic animals may be isolated from their respective sera and eggs and combined in vitro to generate a binding site capable of binding an antigen.

One aspect of the present invention, therefore, concerns transgenic vertebrate animals such as transgenic birds, for example, transgenic chickens, 30 comprising a recombinant nucleic acid molecule and which may (though optionally) expresses a heterologous gene in one or more cells in the animal. Suitable methods for the generation of transgenic animals are known in the art and

are described in, for example, WO 99/19472, published April 22, 1999; WO 00/11151, published March 2, 2000; and WO 00/56932, published September 28, 2000, the disclosures of which are incorporated herein by reference in their entirety.

5 Embodiments of the methods for the production of a heterologous polypeptide by avian tissue such as oviduct tissue and the production of eggs which contain heterologous protein involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, for example, a serine recombinase such as phiC31 integrase, so that the vector can  
10 integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells produced in the previous steps. Deriving a mature transgenic avian from the blastodermal cells optionally involves transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as  
15 the embryo is allowed to develop.

Another alternative may be to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In another embodiment, the cells of a blastodermal embryo are transfected  
20 or transduced with the vector and integrase directly within the embryo. It is contemplated, for example, that the recombinant nucleic acid molecules of the present invention may be introduced into a blastodermal embryo by direct microinjection of the DNA into a stage X or earlier embryo that has been removed from the oviduct. The egg is then returned to the bird for egg white deposition,  
25 shell development and laying. The resulting embryo is allowed to develop and hatch, and the chick allowed to mature.

In one embodiment, a transgenic bird of the present invention is produced by introducing into embryonic cells such as, for instance, isolated avian blastodermal cells, a nucleic acid construct comprising an attB recombination site  
30 capable of recombining with a pseudo-attP recombination site found within the nuclear genome of the organism from which the cell was derived, and a nucleic acid fragment of interest, in a manner such that the nucleic acid fragment of

interest is stably integrated into the nuclear genome of germ line cells of a mature bird and is inherited in normal Mendelian fashion. It is also within the scope of the invention that the targeted cells for receiving the transgene have been engineered to have a heterologous attP recombination site, or other recombination site, integrated 5 into the nuclear genome of the cells, thereby increasing the efficiency of recognition and recombination with a heterologous attB site.

In either case, the transgenic bird produced from the transgenic blastodermal cells is known as a "founder". Some founders can be chimeric or mosaic birds if, for example, microinjection does not deliver nucleic acid 10 molecules to all of the blastodermal cells of an embryo. Some founders will carry the transgene in the tubular gland cells in the magnum of their oviducts and will express the heterologous protein encoded by the transgene in their oviducts. If the heterologous protein contains the appropriate signal sequences, it will be secreted into the lumen of the oviduct and onto the yolk of an egg.

15 Some founders are germ-line founders. A germ-line founder is a founder that carries the transgene in genetic material of its germ-line tissue, and may also carry the transgene in oviduct magnum tubular gland cells that express the heterologous protein. Therefore, in accordance with the invention, the transgenic bird will have tubular gland cells expressing the heterologous protein and the 20 offspring of the transgenic bird will also have oviduct magnum tubular gland cells that express the selected heterologous protein. (Alternatively, the offspring express a phenotype determined by expression of the exogenous gene in a specific tissue of the avian.)

The stably modified oviduct cells will express the heterologous 25 polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

This description uses gene nomenclature accepted by the Cucurbit Genetics 30 Cooperative as it appears in the Cucurbit Genetics Cooperative Report 18:85 (1995), which are incorporated herein by reference in its entirety.

The disclosures of publications, patents, and published patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

5 It will be apparent to those skilled in the art that various modifications, combinations, additions, deletions and variations can be made in the present invention without departing from the scope or spirit of the invention. For instance, features illustrated or described as part of one embodiment can be used in another embodiment to yield a still further embodiment. It is intended that the present  
10 invention covers such modifications, combinations, additions, deletions and variations as come within the scope of the appended claims and their equivalents.

The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patents and patents cited throughout the  
15 present application are hereby incorporated by reference in their entireties.

**Example 1: Phage phiC31 Integrase Functions in Avian Cells.**

(a) A luciferase vector bearing either an attB (SEQ ID NO: 2 shown in Fig. 10) or attP (SEQ ID NO: 3 shown in Fig. 11) site was co-transfected with an integrase expression vector CMV-C31int (SEQ ID NO: 1) into DF-1 cells, a chicken fibroblast cell line. The cells were passaged several times and the luciferase levels were assayed at each passage.

Cells were passaged every 3-4 days and one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a percentage of the expression measured 4 days after transfection. A luciferase expression vector bearing an attP site as a control was also included.

As can be seen in Fig. 2, in the absence of integrase, luciferase expression from a vector bearing attP or attB decreased to very low levels after several days. However, luciferase levels were persistent when the luciferase vector bearing attB  
30 was co-transfected with the integrase expression vector, indicating that the luciferase vector had stably integrated into the avian genome.

(b) A drug-resistance colony formation assay was used to quantitate integration efficiency. The puromycin resistance expression vector pCMV-pur was outfitted with an attB (SEQ ID NO: 4 shown in Fig. 12) or an attP (SEQ ID NO: 5 shown in Fig. 13) sites. Puromycin resistance vectors bearing attB sites were cotransfected 5 with phiC31 integrase or a control vector into DF-1 cells. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

In the absence of co-transfected integrase expression, few DF-1 cell colonies were observed after survival selection. When integrase was co-expressed, 10 multiple DF-1 cell colonies were observed, as shown in Fig. 3. Similar to the luciferase expression experiment, the attB sequence (but not the attP sequence) was able to facilitate integration of the plasmid into the genome. Fig. 3 also shows that phiC31 integrase functions at both 37° Celsius and 41° Celsius. Integrase also functions in quail cells using the puromycin resistance assay, as shown in Fig. 4.

15 (c) The CMV-pur-attB vector (SEQ ID NO: 4) was also cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site (SEQ ID NO: 6 shown in Fig. 14) into DF-1 cells and the phiC31 integrase expression vector CMV-C31int (SEQ ID NO: 1). After puromycin selection for 12 days, the colonies were viewed with UV light to determine the percentage of cells 20 that expressed EGFP. Approximately 20% of puromycin resistant colonies expressed EGFP in all of the cells of the colony, as shown in Fig. 5, indicating that the integrase can mediate multiple integrations per cell.

(d) PhiC31 integrase promoted the integration of large transgenes into avian cells. A puromycin expression cassette comprising a CMV promoter, puromycin 25 resistance gene, polyadenylation sequence and the attB sequence was inserted into a vector containing a 12.0 kb lysozyme promoter and the human interferon  $\alpha$ 2b gene (SEQ ID NO: 7 shown in Fig. 15) and into a vector containing a 10.0 kb ovomucoid promoter and the human interferon  $\alpha$ 2b gene (SEQ ID NO: 8) as shown in Fig. 16.

30 DF-1 cells were transfected with donor plasmids of varying lengths bearing a puromycin resistance gene and an attB sequence in the absence or presence of an integrase expression plasmid. Puromycin was added to the culture media to kill

those cells which did not contain a stably integrated copy of the puromycin resistance gene. Cells with an integrated gene formed colonies in the presence of puromycin in 7-12 days. The colonies were visualized by staining with methylene blue and the entire 60 mm culture dish was imaged.

5 PhiC31 integrase mediated the efficient integration of both vectors as shown in Fig. 7.

#### Example 2: Cell Culture Methods

DF-1 cells were cultured in DMEM with high glucose, 10% fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin and 100 µg/ml streptomycin at 10 37° Celsius and 5% CO<sub>2</sub>. A separate population of DF-1 cells was grown at 41° Celsius. These cells were adapted to the higher temperature for one week before they were used for experiments.

Quail QT6 cells were cultured in F10 medium (Gibco) with 5% newborn 15 calf serum, 1% chicken serum heat inactivated (at 55° Celsius for 45 mins), 10 units/ ml penicillin and 10 µg/ml streptomycin at 37° Celsius and 5% CO<sub>2</sub>.

#### Example 3: Selection and Assay Methods

(a) Puromycin selection assay: About 0.8 x 10<sup>6</sup> DF-1 (chicken) or QT6 (quail) cells 20 were plated in 60 mm dishes. The next day, the cells were transfected as follows:

10 to 50 ng of a donor plasmid and 1 to 10 µg of an Integrase-expressing plasmid DNA were mixed with 150 µl of OptiMEM. 15 µl of DMRIE-C was mixed with 150 µl of OptiMEM in a separate tube, and the mixtures combined and incubated for 15 mins. at room temperature.

25 While the liposome/DNA complexes were forming, the cells were washed with OptiMEM and 2.5 ml of OptiMEM was added. After 15 minutes, 300 µl of the DNA-lipid mixture was added drop wise to the 2.5 ml of OptiMEM covering the cell layers. The cells were incubated for 4-5 hours at either 37° Celsius or 41° Celsius, 5% CO<sub>2</sub>. The transfection mix was replaced with 3 mls of culture media. 30 The next day, puromycin was added to the media at a final concentration of 1 µg/ml, and the media replaced every 2 to 4 days. Puromycin resistant colonies were counted or imaged 10-12 days after the addition of puromycin.

(b) Luciferase assay: Chicken DF-1 or quail QT6 cells ( $0.8 \times 10^6$ ) were plated in 60 mm dishes. Cells were transfected as described above. The cells from a plate were transferred to a new 100 mm plate when the plate became confluent, typically on day 3-4, and re-passaged every 3-4 days.

5 At each time point, one-third of the cells from a plate were replated, and one-third were harvested for the luciferase assay. The cells were pelleted in an eppendorf tube and frozen at  $-70^{\circ}\text{C}$ .

The cell pellet was lysed in 200  $\mu\text{l}$  of lysis buffer (25 mM Tris-acetate, pH7.8, 2mM EDTA, 0.5% Triton X-100, 5% glycerol). Sample (5 $\mu\text{l}$ ) was assayed  
10 using the Promega BrightGlo reagent system.

- 15 (c) Visualization of EGFP: EGFP expression was visualized with an inverted microscope with FITC illumination [Olympus IX70, 100 W mercury lamp, HQ-FITC Band Pass Emission filter cube, exciter 480/40 nm, emission 535/50 nm, 20X phase contrast objective (total magnification was 2.5 x 10 x 20)].
- (d) Staining of cell colonies: After colonies had formed, typically after 7-12 days of culture in puromycin medium, the cells were fixed in 2% formaldehyde, 0.2% glutaraldehyde for 15 mins, and stained in 0.2% methylene blue for 30 mins. followed by several washes with water. The plates were imaged using a standard CCD camera in visible light.

20

**Example 4: Production of Genetically Transformed Avian Cells**

Avian stage X blastodermal cells are used as the cellular vector for the transgenes. Stage X embryos are collected and the cells dispersed and mixed with plasmid DNA. The transgenes are then introduced to blastodermal cells via  
25 electroporation. The cells are immediately injected back into recipient embryos.

The cells are not cultured for any time period to ensure that they remain capable of contributing to the germline of resulting chimeric embryos. However, because there is no culture step, cells that bear the transgene cannot be identified. Typically, only a small percentage of cells introduced to an embryo will bear a  
30 stably integrated transgene (0.01 to 1%). To increase the percentage of cells bearing a transgene, therefore, the transgene vector bears an attB site and is co-electroporated with a vector bearing the CMV promoter driving expression of the

phiC31 transgene (CMV-C31int (SEQ ID NO: 1)). The integrase then drives integration of the transgene vector into the nuclear genome of the avian cell and increases the percentage of cells bearing a stable transgene.

(a) Preparation of avian stage X blastodermal cells:

- 5        i)      Collect fertilized eggs from Barred Rock or White leghorn chickens (*Gallus gallus*) or quail (*Japonica coturnix*) within 48 hrs. of laying;
- ii)     Use 70% ethanol to clean the shells;
- iii)    Crack the shells and open the eggs;
- iv)     Remove egg whites by transferring yolks to opposite halves of shells, repeating to remove most of the egg whites;
- 10       v)     Put egg yolks with embryo discs facing up into a 10cm petri dish;
- vi)     Use an absorbent tissue to gently remove egg white from the embryo discs;
- vii)    Place a Whatman filter paper 1 ring over the embryos;
- 15       viii)   Use scissors to cut the membranes along the outside edge of the paper ring while gently lifting the ring/embryos with a pair of tweezers;
- ix)     Insert the paper ring with the embryos at a 45 degrees angle into a petri dish containing PBS-G solution at room temperature;
- x)     After ten embryo discs are collected, gently wash the yolks from the blastoderm discs using a Pasteur pipette under a stereo microscope;
- 20       xi)   Cut the discs by a hair ring cutter (a short piece of human hair is bent into a small loop and fastened to the narrow end of a Pasteur pipette with Parafilm);
- xii)   Transfer the discs to a 15 ml sterile centrifuge tube on ice;
- 25       xiii)   Place 10 to 15 embryos per tube and allow to settle to the bottom (about 5 mins.);
- xiv)   Aspirate the supernatant from the tube;
- xv)     Add 5 mls of ice-cold PBS without Ca<sup>++</sup> and Mg<sup>++</sup>, and gently pipette 4 to 5 times using a 5 mls pipette;
- 30       xvi)   Incubate in ice for 5-7 mins. to allow the blastoderms to settle, and aspirate the supernatant;

- xvii) Add 3 mls of ice cold 0.05% trypsin/0.02% ETDA to each tube and gently pipette 3 to 5 times using a 5 ml pipette;
- xviii) Put the tube in ice for 5 mins. and then flick the tube by finger 40 times.  
Repeat;
- 5 xix) Add 0.5 mls FBS and 3-5 mls BDC medium to each tube and gently pipette 5-7 times using a 5 ml pipette;
- xx) Spin at 500 rpm (RCF 57 x g) at 4° Celsius for 5 mins;
- xxi) Remove the supernatant and add 2 mls ice cold BDC medium into each tube; and
- 10 xxii) Resuspend the cells by gently pipetting 20-25 times; and
- xxiii) Determine the cell titer by hemacytometer and ensure that about 95% of all BDCs are single cells, and not clumped.
- (b) Transfection of linearized plasmids into blastodermal cells by small scale electroporation:
- 15 i) Centrifuge the blastodermal cell suspension from step (xxiii) above at RCF 57 x g, 4° Celsius, for 5 mins;
- ii) Resuspend cells to a density of 1-3 x 10<sup>6</sup> per ml with PBS without Ca<sup>2+</sup> and Mg<sup>2+</sup>;
- 20 iii) Add linearized DNA, 1-30 µg per 1-3 x 10<sup>5</sup> blastodermal cells in an eppendorf tube at room temperature. Add equimolar molar amounts of the non-linearized transgene plasmid bearing an attB site, and an integrase expression plasmid;
- iv) Incubate at room temperature for 10 mins;
- v) Aliquot 100 µl of the DNA-cell mixture to a 0.1 cm cuvette at room 25 temperature;
- vi) Electroporate at 240 V and 25 µFD (or 100 V and 125 µFD for quail cells) using, for example, a Gene Pulser II™ (BIO-RAD).
- vii) Incubate the cuvette at room temperature for 1-10 mins.
- viii) Before the electroporated cells are injected into a recipient embryo, they 30 are transferred to a eppendorf tube at room temperature. The cuvette is washed with 350 µl of media, which is transferred to the eppendorf, spun at room temperature and re-suspended in 0.01-0.3 ml medium;

- ix) Inject 1-10  $\mu$ l of cell suspension into the subgerminal cavity of an non-irradiated or, for example, an irradiated (e.g., with 300-900 rads) stage X egg. Shell and shell membrane are removed and, after injection, resealed according to U.S. Patent No. 6,397,777, issued June 6, 2002, the disclosure of which is incorporated herein by reference in its entirety; and
- 5           x) The egg is then incubated to hatching.
- (c) Blastodermal Cell Culture Medium:
- 10           i) 409.5 mls DMEM with high glucose, L-glutamine, sodium pyruvate, pyridoxine hydrochloride;
- ii) 5 mls Men non-essential amino acids solution, 10 mM;
- iii) 5 mls Penicillin-streptomycin 5000 U/ml each;
- iv) 5 mls L-glutamine, 200 mM;
- v) 75 mls fetal bovine serum; and
- 15           vi) 0.5 mls  $\beta$ -mercaptoethanol, 11.2mM.

**Example 5: Transfection of Stage X Embryos with attB Plasmids**

(a) DNA-PEI: Twenty-five  $\mu$ g of a phage phiC31 integrase expression plasmid (pCMV-int), and 25  $\mu$ g of a luciferase-expressing plasmid (p $\beta$ -actin-GFP-attB) are combined in 200  $\mu$ l of 28 mM Hepes (pH 7.4). The DNA/Hepes is mixed with an equal volume of PEI which has been diluted 10-fold with water. The DNA/Hepes/PEI is incubated at room temperature for 15 mins Three to seven  $\mu$ l of the complex are injected into the subgerminal cavity of windowed stage X white leghorn eggs which are then sealed and incubated as described in U.S. Patents No. 20 6,397,777, issued June 6, 2002. The complexes will also be incubated with blastodermal cells isolated from stage X embryos which are subsequently injected into the subgerminal cavity of windowed irradiated stage X white leghorn eggs. 25 Injected eggs are sealed and incubated as described above.

25           (b) Adenovirus-PEI:  
30           Two  $\mu$ g of a phage phiC31 integrase expression plasmid (pCMV-int), 2  $\mu$ g of a GFP expressing plasmid (p $\beta$ -actin-GFP-attB) and 2  $\mu$ g of a luciferase expressing plasmid (pGLB) were incubated with 1.2  $\mu$ l of JetPEI<sup>TM</sup> in 50  $\mu$ l of 20 mM Hepes

buffer (pH7.4). After 10 mins at 25°C, 3 x 10<sup>9</sup> adenovirus particles (Ad5-Null, Qbiogene) were added and the incubation continued for an additional 10 mins. Embryos are transfected in ovo or ex ovo as described above.

5

#### Example 6: Stage I Cytoplasmic Injection

Production of transgenic chickens by cytoplasmic DNA injection using DNA injection directly into the germinal disk as described in Sang et al, Mol. Reprod. Dev., 1: 98-106 (1989); Love et al, Biotechnology, 12: 60-63 (1994) incorporated herein by reference in their entireties.

10

In the method of the present invention, fertilized ova, or stage I embryos, are isolated from euthanized hens 45 mins. to 4 hrs. after oviposition of the previous egg. Alternatively, eggs were isolated from hens whose oviducts have been fistulated according to the techniques of Gilbert & Wood-Gush, J. Reprod. Fertil., 5: 451-453 (1963) and Pancer et al, Br. Poult. Sci., 30: 953-7 (1989) incorporated herein in their entireties.

15

An isolated ovum was placed in dish with the germinal disk upwards. Ringer's buffer medium was then added to prevent drying of the ovum. Any suitable microinjection assembly and methods for microinjecting and reimplanting avian eggs are useful in the method of cytoplasmic injection of the present

20

invention. A particularly suitable apparatus and method for use in the present invention is described in U.S. Patent Application Serial No: 09/919,143, published July 31, 2001, the disclosure of which is incorporated in its entirety herein by reference. The avian microinjection system described in the '143 Application allowed the loading of a DNA solution into a micropipette, followed by prompt

25

positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk. Injected embryos could then be surgically transferred to a recipient hen as described, for example, in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J. Reprod. Fertil., 100: 447-449 (1994).

30

The embryo was allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which was incubated until hatching of the chick. Injected embryos were surgically transferred to recipient hens via the ovum transfer method

of Christmann et al in PCT/US01/26723, published August 27, 2001, the disclosure of which is incorporated herein by reference in its entirety, and hard shell eggs were incubated and hatched.

5        Approximately 25 nl of DNA solution (about 60ng/μl) with either integrase mRNA or protein were injected into a germinal disc of stage I White Leghorn embryos obtained 90 minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was 100 ng/μl, and the concentration of integrase protein was 66 ng/μl.

10      To synthesize the integrase mRNA, a plasmid template encoding the integrase protein was linearized at the 3' end of the transcription unit. mRNA was synthesized, capped and a polyadenine tract added using the mMESSAGE mMACHINE T7 Ultra Kit™ (Ambion, Austin, TX). The mRNA was purified by extraction with phenol and chloroform and precipitated with isopropanol. The integrase protein was expressed in *E. coli* and purified as described by Thorpe et al,  
15      Mol. Microbiol., 38: 232-241 (2000).

A plasmid encoding for the integrase protein is transfected into the target cells. However, since the early avian embryo transcriptionally silent until it reaches about 22,000 cells, injection of the integrase mRNA or protein was expected to result in better rates of transgenesis, as shown in the Table 1 below.

20      The chicks produced by this procedure were screened for the presence of the injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002).

Table 1: Summary of cytoplasmic injection results using different integrase strategies

Experimental group	Ovum transfers	Hard shells produced (%)	Chicks hatched (%) *	Transgenic chicks (%) ‡
No Integrase	5164	3634 (70%)	500 (14%)	58 (11.6%)
Integrase mRNA	1109	833 (75%)	115 (13.8%)	19 (16.5%)
Integrase protein	374	264 (70.6%)	47 (17.8%)	16 (34%)

\* : Percentages based on the number of hard shells

‡ : Percentages based on the number of hatched birds

5

**Example 7: Characterization of phiC31 Integrase-Mediated Integration Sites in the Chicken Genome**

To characterize phiC31-mediated integration into the chicken genome, a plasmid rescue method was used to isolate integrated plasmids from transfected and selected chicken fibroblasts. Plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10, shown in Fig. 18) does not have BamH I or Bgl II restriction sites. Genomic DNA from cells transformed with pCR-XL-TOPO-CMV-pur-attB was cut with BamH I or Bgl II (either or both of which would cut in the flanking genomic regions) and religated so that the genomic DNA surrounding the integrated plasmid would be captured into the circularized plasmid. The flanking DNA of a number of plasmids were then sequenced.

DF-1 cells (chicken fibroblasts),  $4 \times 10^5$  were transfected with 50 ng of pCR-XL-TOPO-CMV-pur-attB and 1  $\mu$ g of pCMV-int. The following day, the culture medium was replaced with fresh media supplemented with 1  $\mu$ g/ml puromycin. After 10 days of selection, several hundred puromycin-resistant colonies were evident. These were harvested by trypsinization, pooled, replated on 10 cm plates and grown to confluence. DNA was then extracted.

Isolated DNA was digested with BamH I and Bgl II for 2-3 hrs, extracted with phenol:chloroform:isoamyl alcohol chloroform:isoamyl alcohol and ethanol precipitated. T4 DNA ligase was added and the reaction incubated for 1 hr at room

temperature, extracted with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol, and precipitated with ethanol. 5 µl of the DNA suspended in 10µl of water was electroporated into 25 µl of Genehogs™ (Invitrogen) in an 0.1 cm cuvette using a GenePulser II (Biorad) set at 1.6 kV, 100 5 ohms, 25 uF and plated on Luria Broth (LB) plates with 5 µg/ml phleomycin (or 25 µg/ml zeocin) and 20 µg/ml kanamycin. Approximately 100 individual colonies were cultured, the plasmids extracted by standard miniprep techniques and digested with Xba I to identify clones with unique restriction fragments.

Thirty two plasmids were sequenced with the primer attB-for (5'-  
10 TACCGTCGACGATGTAGGTCACGGTC-3') (SEQ ID NO: 12) which allows sequencing across the crossover site of attB and into the flanking genomic sequence. All of plasmids sequenced had novel sequences inserted into the crossover site of attB, indicating that the clones were derived from plasmid that had integrated into the chicken genome via phiC31 integrase-mediated  
15 recombination.

The sequences were compared with sequences at GenBank using Basic Local Alignment Search Tool (BLAST). Most of the clones harbored sequences homologous to Gallus genomic sequences in the TRACE database.

20 **Example 8: Insertion of a Wild-Type attP Site into the Avian Genome**

**Augments Integrase-Mediated Integration and Transgenesis**

The chicken B-cell line DT40 cells (Buerstedde et al, E.M.B.O. J., 9: 921-  
927 (1990)) are useful for studying DNA integration and recombination processes  
(Buerstedde & Takeda, Cell, 67:179-88 (1991)). DT40 cells were engineered to  
25 harbor a wild-type attP site isolated from the Streptomyces phage phiC31. Two independent cell lines were created by transfection of a linearized plasmid bearing an attP site linked to a CMV promoter driving the resistance gene to G418 (DT40-NLB-attP) or bearing an attP site linked to a CMV promoter driving the resistance gene for puromycin (DT40-pur-attP). The transfected cells were cultured in the  
30 presence of G418 or puromycin to enrich for cells bearing an attP sequence stably integrated into the genome.

A super-coiled luciferase vector bearing an attB (SEQ ID NO: 2 shown in Fig. 10) was co-transfected, together with an integrase expression vector CMV-C31int (SEQ ID NO: 1) or a control, non-integrase expressing vector (CMV-BL) into wild-type DT40 cells and the stably transformed lines DT40-NLB-attP and 5 DT40-pur-attP.

Cells were passaged at 5, 7 and 14 days post-transfection and about one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a percentage of the expression measured 5 days after transfection. As can be seen in Fig. 21, in the absence of integrase, or in the 10 presence of integrase but in the DT40 cells lacking an inserted wild-type attP site, luciferase expression from a vector bearing attB progressively decreased to very low levels. However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector into the attP bearing cell lines DT40-NLB-attP and DT40-pur-attP. Inclusion of an attP 15 sequence in the avian genome augments the level of integration efficiency beyond that afforded by the utilization of endogenous pseudo-attP sites.

**Example 9: Generation of attP Transgenic Cell Line  
and Birds Using an NLB Vector**

- 20       The NLB-attP retroviral vector is injected into stage X chicken embryos laid by pathogen-free hens. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane is cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10 µl of concentrated retrovirus, approximately  $2.5 \times 10^5$  IU, is injected into the subgerminal cavity of the embryo.
- 25       The egg shell is resealed with a hot glue gun. Suitable methods for the manipulation of avian eggs, including opening and resealing hard shell eggs are described in U.S. Patent Serial Nos: 5,897,998, issued May 27, 1999 and 6,397,777, issued June 4, 2002, the disclosures of which are herein incorporated by reference in their entireties.
- 30       Typically, 25% of embryos hatch 21 days later. The chicks are raised to sexual maturity and semen samples are taken. Birds that have a significant level of the transgene in sperm DNA will be identified, typically by a PCR-based assay.

Ten to 25% of the hatched roosters will be able to give rise to G1 transgenic offspring, 1 to 20% of which may be transgenic. DNA extracted from the blood of G1 offspring is analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of transgenic roosters, each with a unique site of 5 attP integration, are then bred to non-transgenic hens, giving 50% of G2 transgenic offspring. Transgenic G2 hens and roosters from the same line can be bred to produce G3 offspring homozygous for the transgene. Homozygous offspring will be distinguished from hemizygous offspring by quantitative PCR. The same procedure can be used to integrate an attB or attP site into transgenic birds.

10

**Example 10: Expression of Immunoglobulin Chain Polypeptides by  
Transgenic Chickens**

Bacterial artificial chromosomes (BACs) containing a 70 kb segment of the chicken ovomucoid gene with the light and heavy chain cDNAs for a human monoclonal antibody inserted along with an internal ribosome entry site into the 3' untranslated region of the ovomucoid gene were equipped with the attB sequence. 15 The heavy and light chain cDNAs were inserted into separate ovomucoid BACs such that expression of an intact monoclonal antibody requires the presence of both BACs in the nucleus.

20 Several hens produced by coinjection of the attB-bearing ovomucoid BACs and integrase-encoding mRNA into stage I embryos produced intact monoclonal antibodies in their egg white. One hen, which had a high level of the light chain ovomucoid BAC in her blood DNA as determined by quantitative PCR particularly expressed the light chain portion of the monoclonal antibody in the egg white at a 25 concentration of 350 nanograms per ml, or approximately 12 µg per egg.

**Example 11: Stage I Cytoplasmic Injection with Integrase Activity and PEI**

Production of transgenic chickens by cytoplasmic DNA injection directly into the germinal disk was done as described in Example 6.

30 DNA (about 60ng/µl) which includes a transgene was placed in approximately 25 nl of aqueous solution with integrase mRNA or integrase protein and was mixed with an equal volume of PEI that had been diluted ten fold. The

mixture was injected into a germinal disc of stage I White Leghorn embryos obtained about 90 minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was about 100 ng/μl, and the concentration of integrase protein was about 66 ng/μl. The integrase mRNA was synthesized  
5 according to Example 6.

Transgenic chicks produced by this procedure using: integrase mRNA/PEI and integrase protein/PEI showed positive results for the presence of heterologously expressed protein in the blood, semen and egg white.

10 **Example 12: Stage I Cytoplasmic Injection with Integrase Activity and NLS**

Production of transgenic chickens by cytoplasmic DNA injection directly into the germinal disk was done as described in Example 6.

DNA which includes a transgene was suspended in 0.25 M KCl and SV40 T antigen nuclear localization signal peptide (NLS peptide, amino acid sequence 15 CGGPKKKRKVG (SEQ ID NO: 13)) was added to achieve a peptide DNA molar ratio of 100:1. The DNA (about 60ng/μl) was allowed to associate with the SV40 T antigen NLS peptide by incubating at 25 degrees C for about 15 minutes.

Integrase mRNA or integrase protein was added to approximately 25 nl of an aqueous DNA/NLS solution, typically, to produce a final concentration of 20 integrase mRNA of about 50 ng/μl, or an integrase protein concentration of about 33 ng/μl. The mixture was injected into a germinal disc of stage I White Leghorn embryos obtained about 90 minutes after oviposition of the preceding egg. The integrase mRNA was synthesized as according to Example 6.

Transgenic chicks produced by this procedure using: integrase mRNA/NLS 25 and integrase protein/NLS showed positive results for the presence of heterologously expressed protein in blood, semen and egg white.

**Example 13: Dispersing of Plasmid DNA in Avian Stage I Embryos**

DNA samples are Cy3 labeled with a Cy3 ULS labeling kit (Amersham 30 Pharmacia Biotech). Briefly, plasmid DNA (1 μg) was sheared to approximately 100 to 500 bp fragments by sonication. Resulting DNA was incubated at 65°C for 15 min in Cy3 ULS labeling solution and unincorporated Cy3 dye was removed by

spin column chromatography (CentriSep, Princeton Separations). The distribution of the DNA in stage I avian embryos was visualized after introduction into the stage I avian embryo. Enough high molecular weight or low molecular weight PEI was added to the DNA to coat the DNA. Typically, PEI was added to the DNA to 5 a concentration of about 5%. Any useful volume of DNA/PEI can be used, for example about 25 nl.

Figure 22 shows an avian stage one embryo containing Cy3 labeled naked DNA. In Figure 22 it can be seen that the DNA is localized to certain areas of the embryo. Figure 23 and Figure 24 show an avian stage one embryo containing Cy3 10 labeled DNA coated with low molecular (22 kD) weight PEI (Figure 23) and high molecular weight (25 kD) PEI (Figure 24). In Figures 23 and 24, it can be seen that the DNA is dispersed throughout the embryos.

These experiments show that DNA/PEI conjugates are distributed more uniformly in the cytoplasm of injected embryos when compared with naked DNA.

15

**Example 14: Production of an attP Transgenic Chicken**

G0 transgenic chickens have been produced as described in Example 9. Several hundred stage X White Leghorn eggs were injected with the NLB-attP vector and about 50 chicks hatched. Sperm from approximately 30% of the 20 hatched roosters has been shown to be positive for the attP site. These hemizygotic chickens are used to generate transgenic G2 chickens homozygotic for the attP site.

**Example 15: Cytoplasmic Injection of attP Stage I Embryos with OMC24-attB-IRES-CTLA4**

Transgenic chickens are produced by cytoplasmic DNA injection directly 25 into the germinal disk of eggs laid by transgenic homozygous attP chickens and fertilized with sperm from the same line of homozygous attP roosters, the line produced as described in Example 14. The cytoplasmic injections are carried out as described in U.S. Patent Application Serial No. 09/919,143, filed July 31, 2001, 30 ('143 Application) and U.S. Patent Application Serial No. 10/251,364, filed September 18, 2002. The disclosures of each of these two patent applications are incorporated herein by reference in their entirety.

Stage I embryos are isolated 45 mins. to 4 hrs. after oviposition of the previous egg. An isolated embryo is placed in a dish with the germinal disk upwards. Ringer's buffer medium is added to prevent drying of the ovum. The avian microinjection system described in the '143 Application allows for the 5 loading of DNA solution into a micropipette, followed by prompt positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk.

Approximately 25 nl of a DNA solution (about 60ng/ $\mu$ l) of the 77 kb OMC24-attB-IRES-CTLA4, disclosed in US Patent Application No. 10/856,218, 10 filed May 28, 2004, the disclosure of which is incorporated in its entirety herein by reference, with either integrase mRNA or protein are injected into a germinal disc of the isolated stage I embryos. Typically, the concentration of integrase mRNA used is 100 ng/ $\mu$ l or the concentration of integrase protein is 66 ng/ $\mu$ l.

To synthesize the integrase mRNA, a plasmid template encoding the 15 integrase protein is linearized at the 3' end of the transcription unit. mRNA is synthesized, capped and a polyadenine tract added using the mMESSAGE mMACHINE T7 Ultra Kit<sup>TM</sup> (Ambion, Austin, TX). The mRNA is purified by extraction with phenol and chloroform and precipitated with isopropanol. The integrase protein is expressed in E. coli and purified as described by Thorpe et al, 20 Mol. Microbiol., 38: 232-241 (2000).

Injected embryos are surgically transferred to a recipient hen as described in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J. Reprod. Fertil., 100: 447-449 (1994). The embryo is allowed to proceed through the natural 25 in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which is incubated until hatching of the chick. Injected embryos are surgically transferred to recipient hens via the ovum transfer method of Christmann et al in PCT/US01/26723, published August 27, 2001, the disclosure of which is incorporated by reference in its entirety, and hard shell eggs are incubated and hatched.

The chicks produced by this procedure are screened for the presence of the 30 injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002). Approximately

20% of the chicks are positive for the transgene. Eggs from each of the mature hens carrying the transgene are positive for CTLA4.

**Example 16: Cytoplasmic Injection of attP Stage I Chicken Embryos with OM10-attB-CTLA4**

5

Transgenic chickens are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by transgenic homozygous attP chickens and fertilized with sperm from the same line of homozygous attP roosters essentially as described in Example 15.

10

Approximately 25 nl of a 60ng/ $\mu$ l DNA solution of the OMC24-attB-IRES-CTLA4 construct of Example 15 with the OMC24 70 kb ovomucoid gene expression controlling region and IRES of the construct replaced with the 10 kb ovomucoid gene expression controlling region of pBS-OVMUP-10, also disclosed in US Patent Application No. 10/856,218, filed May 28, 2004, is injected into a fertilized germinal disc of stage I embryos along with an integrase protein. The concentration of integrase protein used is 66 ng/ $\mu$ l.

15

Injected embryos are then surgically transferred to a recipient hen, hard shell eggs are produced, incubated and hatched. Approximately 30% of the chicks are positive for the transgene. Eggs from each of the mature hens carrying the transgene are positive for CTLA4.

20

**Example 17: Production of attP Transgenic Quail Using an NLB vector**

25

The NLB-attP retroviral vector is injected into stage X quail embryos laid by pathogen-free quail. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10  $\mu$ l of concentrated retrovirus, approximately  $1.0 \times 10^5$  IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun.

30

Typically, 25% of embryos hatch. The chicks are raised to sexual maturity and semen samples are taken. Birds that have a significant level of the transgene in their sperm DNA will be identified, typically by a PCR-based assay. Of the

hatched G0 male quail, about 1% to about 20% are transgenic. The transgenic G0 quail are bred to nontransgenic quail to produce hemizygotic G1 offspring. DNA extracted from the blood of G1 offspring is analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of hemizygotic 5 transgenic male quail, each with a unique site of attP integration, are then bred to non-transgenic quail giving G2 offspring, 50% of which are transgenic. Transgenic G2 male and female from the same line are then bred to produce G3 offspring homozygous for the transgene. Homozygous offspring are distinguished from hemizygous offspring by quantitative PCR.

10

**Example 18: Cytoplasmic Injection of attP Stage I Quail Embryos with OMC24-attB-IRES-G-CSF**

Transgenic quail are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by fully transgenic homozygous attP quail produced 15 as described in Example 17. The cytoplasmic injections are carried out essentially as described in the '143 Application and U.S. Patent Application Serial No. 10/251,364, filed September 18, 2002.

Stage I embryos from homozygous attP quail fertilized with sperm from a homozygous attP quail are isolated approximately 90 minutes after oviposition of 20 the previous egg. An isolated embryo is placed in a dish with the germinal disk upwards. Ringer's buffer medium is added to prevent drying of the ovum. The avian microinjection system described in the '143 Application is used to inject approximately 25 nl of a DNA solution (about 60ng/ $\mu$ l) of OMC24-attB-IRES-CTLA4, with the CTLA coding sequence replaced with the coding sequence for a 25 human-granulocyte colony stimulating factor, and integrase protein into the germinal disc of the stage I quail embryos. The concentration of integrase protein used is 66 ng/ $\mu$ l.

Injected embryos are surgically transferred to a recipient quail essentially as described in Olsen & Neher, J. Exp. Zool., 109: 355-66 (1948) and Tanaka et al, J. 30 Reprod. Fertil., 100: 447-449 (1994). The embryo is allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The

transgenic embryo is then laid as a hard-shell egg which is incubated until hatching of the chick.

The chicks produced by this procedure are screened for the presence of the injected transgene using a high throughput PCR-based screening procedure as described in Harvey et al, Nature Biotech., 20: 396-399 (2002). Approximately 20% of the chicks are positive for the transgene. Eggs from each of the mature female quail carrying the transgene are positive for G-CSF.

**Example 19: Generation of attP Transgenic Duck**

10                   **Using an NLB vector**

The NLB-attP retroviral vector is injected into stage X Duck embryos laid by pathogen-free Ducks. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. About 1 to 10  $\mu$ l of concentrated retrovirus, approximately  $2.5 \times 10^5$  IU, is injected into the 15 subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun

Homozygous G3 offspring are obtained essentially as described in Example 17 for quail.

**Example 20: Stage I Cytoplasmic Injection of attP Stage I Duck Embryos with**  
20                   **OMC24-attB-IRES-CTLA4**

Transgenic ducks are produced by cytoplasmic DNA injection directly into the germinal disk of eggs laid by homozygous attP ducks fertilized with sperm from homozygous attP ducks. The injection of the stage I embryos is carried out essentially as described in the '143 Application and U.S. Patent Application Serial 25 No. 10/251,364, filed September 18, 2002. Approximately 25 nl of a DNA solution (about 60ng/ $\mu$ l) of OMC24-attB-IRES-CTLA4, with the CTLA4 coding region replaced with a coding sequence for human erythropoietin, and integrase encoding mRNA and protein is injected into the germinal disc of the stage I embryos. The concentration of integrase mRNA used is 100 ng/ $\mu$ l. The injected 30 embryos are surgically transferred to a recipient duck and the embryo is allowed to proceed through the natural in vivo cycle of albumin deposition and hard-shell formation. The transgenic embryo is laid as a hard-shell egg which is incubated

until hatching and the chicks are screened for the presence of the injected transgene. Approximately 20% of the chicks are positive for the transgene. Eggs from each of the mature female ducks carrying the transgene are positive for erythropoietin.

5

**Example 21: Production of Transchromosomal Chickens Using Satellite DNA-Based Artificial Chromosomes**

Satellite DNA-based artificial chromosomes (ACEs, as described in Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172) were isolated by a 10 dual laser high-speed flow cytometer as described previously (de Jong, G, et al. Cytometry 35: 129-133, 1999).

The flow-sorted chromosomes were pelleted by centrifugation of a 750 $\mu$ l sample containing approximately  $10^6$  chromosomes at 2500 x g for 30 min at 4°C. The supernatant, except the bottom 30 microliters ( $\mu$ l) containing the 15 chromosomes, was removed resulting in a concentration of about 7000 to 11,500 chromosomes per  $\mu$ l of injection buffer (Monteith, et al. Methods Mol Biol 240: 227-242, 2004). Depending on the number of chromosomes to be injected, 25-100 nanoliters (nl) of injection buffer was injected per embryo.

Embryos for this study were collected from 24-36 week-old hens from 20 commercial White Leghorn variety of *G. gallus*. Embryo donor hens were inseminated weekly using pooled semen from roosters of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens were euthanized 2h post oviposition by cervical dislocation. Typically, oviposition is followed by ovulation 25 of the next egg after around 24 minutes (Morris, Poultry Science 52: 423-445, 1973). The recently ovulated and fertilized eggs were collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium (Tanaka, et al. J Reprod Fertil 100: 447-449, 1994) and maintained at 41°C until microinjection.

30 Cytoplasmic injection of artificial chromosomes was achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143, filed July 31, 2001. Chromosomes were injected into the Stage I embryos at a single

site. Each embryo was cytoplasmically injected with approximately: 175, 250, 350, 450, 550, 800 or >1000 chromosomes. The chromosomes were injected in a suspension of 25-100 nanoliters (nl) of injection buffer.

Following microinjection, the embryos were transferred to the oviduct of 5 recipient hens using an optimized ovum transfer (OT) procedure (Olsen, M and Neher, B. J Exp Zool 109: 355-66, 1948), with the exception that the hens were anesthetized by Isofluorane gas. Typically, about 26h after OT, the recipient hens lay a hard shell egg containing the manipulated ovum. Eggs were incubated for 21 days in a regular incubator until hatching of the birds.

10 The chromosomes were injected into the embryos over a 9 day period. The chromosomes were divided into three batches for delivery to the embryos each batch being injected over a three day period. Chromosomes were introduced into the embryos by a single injection using the microinjection assembly disclosed in the '143 patent application. Following injection, each egg was transferred to a 15 recipient hen. A total of 301 transfers were performed, resulting in 226 (75%) hard shells and 87 hatched chicks (38%, see Table 2).

**Table 2: Hatching of embryos microinjected with satellite DNA-based artificial chromosomes.**

20

	Ovum transfers	Hard shells produced	hatched birds
1 <sup>st</sup> batch	71	53	15
2 <sup>nd</sup> batch	113	80	33
3 <sup>rd</sup> batch	117	93	39
<b>Totals</b>	<b>301</b>	<b>226 (75%)</b>	<b>87 (38%)</b>

25

Previous experiments have determined that hatching is not significantly affected when embryos were injected with up to 100nl of injection buffer. Satellite DNA-based artificial chromosomes were injected in suspensions of between 25-100nl of injection buffer.

As discussed, the embryos were injected with one of seven different numbers of artificial chromosomes. There was shown to be a correlation between the number of chromosomes injected per egg and the hatch rate. All transchromosomal birds in the present study were obtained from embryos injected

with 550 chromosomes or less (see Table 3). There was no significant difference in the hatching rates observed between the experimental groups (batches 1, 2 and 3).

Six transchromosomal founders were produced based on two separate PCR analysis (6.8%, see Table 3) using primers which anneal to the puromycin resistance gene (about 75 copies of the pur<sup>R</sup> gene are present on the chromosome. All positive birds appear normal.

10 **Table 3: Effect of the number of Chromosomes injected per embryo on hatching and number of transchromosomal birds produced.**

# chromosomes injected per embryo	# of hard shells	# chicks hatched	# of positive birds (bird tag #)
175	31	11 (35%)	3 (BB7478, BB7483, BB7515)
250	51	25 (49%)	1 (BB 7499)
350	15	6 (40%)	0
450	31	11 (35%)	0
550	39	17 (43%)	2 (BB7477, BB7523)
800	26	5 (19%) *	0
1000	33	10 (30%) *	0
<b>Totals</b>	<b>226</b>	<b>87 (38%)</b>	<b>6 (6.8%)</b>

\*: hatching rates of embryos injected with >550 chromosomes was significantly lower (p < 0.025)

15 To confirm the PCR results, erythrocytes from all PCR-positive birds as well as fibroblast cells derived from skin biopsies of 5 PCR-positive birds were analyzed by interphase and metaphase FISH using a mouse-specific major satellite DNA probe (Co, et al. Chromosome Res 8: 183-191, 2000). Five of the six chicks (5.3% out of total number of chicks analyzed) tested by FISH were positive in at least one cell type (see Table 4) at 3 weeks of age. FISH analysis of erythrocytes was repeated when the birds reached 8 weeks of age and had tripled their body weight. Similar numbers of artificial chromosome-positive cells found in each bird were observed in this second FISH analysis.

**Table 4: Summary of FISH analysis of Red Blood Cells (RBCs) and fibroblast cells derived from transchromosomal birds. Fibroblast cells from hen # 7515 were not available for analysis.**

Bird #	Sex of Bird	% of artificial chromosome positive RBCs by FISH	% of artificial chromosome positive fibroblasts by FISH
BB7499	Female	77%	87%
BB7483	Female	0.8%	0%
BB7477	Male	3%	2.8%
BB7478	Male	15%	3%
BB7515	Female	1.3%	NA
BB7523	Male	0%	0%
Neg. control	-	0%	0%

5

To verify the chromosomes were intact, metaphase spreads from fibroblast cells derived from founders were made as described previously (Garside and Hillman (1985) Experientia 41: 1183-1184). FISH analysis of metaphase spreads using the major satellite DNA probe showed the artificial chromosomes appear intact, with no apparent fragmentation or translocation onto the chicken's chromosomes. FISH analysis using a mouse minor satellite probe, which detects the centromeric region of the introduced chromosomes (Wong and Rattner (1988) J. Nucleic Acids Res 16: 11645-11661), demonstrated the centromere of the chromosomes was intact. Furthermore, the percentage of satellite DNA-based artificial chromosomes -positive cells from metaphase spreads agreed closely to those observed in interphase FISH.

Analysis of G1 embryos from test bird BB7499 has shown the artificial chromosome to be transmitted through the germline. In addition, sperm from BB7499 was shown to test positive for the artificial chromosome which will also provide for germline transmission of the artificial chromosome.

**Example 22: Production of EPO and G-CSF Vectors for the Production of Transchromosomal Chickens**

Two vectors were constructed for introduction into Satellite DNA-based artificial chromosomes. 1OMC24-IRES1-EPO - ChromattB was constructed by inserting an EPO coding sequence into an OMC24-IRES BAC clone disclosed in

US Patent Application No. 10/856,218, filed May 28, 2004, the disclosure of which is incorporated in its entirety herein by reference. The EPO coding sequence was inserted in the clone so as to be under the control of the ovomucoid promoter. That is, the EPO coding sequence was inserted in place of the LC portion of OMC-  
5 IRES-LC. An attB site and a hygromycin<sup>R</sup> coding sequence were also inserted into the vector in such a manner as to facilitate recombination into an attP site in a SATAC artificial chromosome (i.e., ACE), as see in FIG. 25. The attP site in the SATAC is located adjacent to an SV40 promoter which provides for expression of the hygromycin<sup>R</sup> coding sequence upon integration of the vector into the attP site  
10 allowing for selection of cells containing a recombinant artificial chromosome (see, for example, US Patent No. 6,743,967, issued June 1, 2004; US Patent No. 6,025,155, issued February 15, 2000 and Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172 (see FIG. 25), the disclosure of each of these two patents and the publication are incorporated in their entirety herein by reference).

15 A coding sequence for G-CSF, which was codon optimized for expression in chicken tubular gland cells, was inserted in the 1OMC24-IRES1-EPO - ChromattB construct in place of the EPO coding sequence to produce 1OMC24-IRES-GCSF - ChrommattB.

20 **Example 23: Production of Erythropoietin and G-CSF Using Artificial Chromosomes in Chickens**

Cells containing the recombinant artificial chromosome are produced and identified as described in Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172. Briefly, 2.5 µg of 1OMC24-IRES1-EPO ChromattB and 2.5 µg of an expression vector which contains a lambda integrase gene (int) having a codon mutation at position 174 to substitute a lysine for a glutamine (pCXLamROK, see Lindenbaum et al Nucleic Acids Res (2004) vol 32 no. 21 e172) are transfected by standard lipofection methodologies into LMTK- cells which contain the platform SATAC (ACE) (A of FIG. 25). Hygromycin resistant cells clones are identified by  
25 standard antibiotic selection methodologies.  
30

Recombinant chromosomes are prepared from the cells and isolated by flow cytometry. The substantially purified artificial chromosomes are introduced into

chickens by microinjection into stage one embryos as disclosed in US Patent Application Nos. 10/679,034, filed October 2, 2003 and 09/919,143, filed July 31, 2001. Resulting chimeric germline transchromosomal avians can be identified by any useful method such as Southern blot analysis.

5

**Example 24: Production of a Monoclonal Antibody Using Drosophila Artificial Chromosomes in Turkey**

Artificial chromosomes comprising a Drosophila chromosome centromere (DAC) are prepared essentially using methods described in US Patent No. 10 6,025,155, issued February 15, 2000, the disclosure of which is incorporated in its entirety herein by reference.

An attB site and a hygromycin<sup>R</sup> coding sequence are inserted into the OMC24-IRES-LC and OMC24-IRES-HC vectors disclosed in US Patent Application No. 10/856,218, filed July 31, 2001, the disclosure of which is incorporated in its entirety herein by reference, which are then each cloned into a DAC essentially as described in Examples 22 and 23. The recombinant DACs are prepared and then isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted to a concentration of about 7000-12,000 chromosomes per  $\mu$ l of injection buffer. 20 Approximately 50 nanoliters (nl) of injection buffer is injected per turkey embryo.

Embryos for this study are collected from actively laying commercial turkeys. Embryo donor turkeys are inseminated weekly using pooled semen from male turkeys of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens are euthanized 2h post oviposition 25 by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 40°C until microinjection.

Cytoplasmic injection of artificial chromosomes containing the OMC24- 30 IRES-LC is achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143. Approximately 500 chromosomes are injected into the Stage I embryos at a single site.

Following microinjection, the embryos are transferred to the oviduct of recipient turkeys essentially as described in Olsen et al, B. J Exp Zool 109: 355-66, 1948. Typically, about one day after OT, the recipient turkeys lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until 5 hatching of the birds.

G2 transchromosomal turkeys are obtained which contain the artificial chromosome in their genome. The artificial chromosome containing the OMC24-IRES-HC is introduced into embryos obtained from the G2 turkeys in essentially the same manner as described for the OMC24-IRES-LC.

10       Eggs from G1 transchromosomal turkeys which contain both the OMC-IRES-LC and OMC24-IRES-HC containing chromosomes in their genome are tested for the presence of intact functional monoclonal antibody. A Costar flat 96-well plate is coated with 100 µl of C Goat-anti-Human kappa at a concentration of 5 µg/ml in PBS. The plate is incubated at 37 °C for two hours. 200 µl of 5% PBA 15 is added to the wells followed by an incubation at 37 °C for about 60-90 minutes followed by a wash. 100 µl of egg white samples (diluted in 1% PBA:LBP) is added to each well and the plate is incubated at 37 °C for about 60-90 min followed by a wash. 100 µl of a 1:2000 dilution of F'2 Goat anti-Human IgG Fc-AP in 1% PBA is added to the wells and the plate is incubated at 37 °C for 60-90 20 min followed by a wash. The antibody is detected by placing 75 µl of 1mg/ml PNPP (p-nitrophenyl phosphate) in 5x developing buffer in each well and incubating for about 10-30 mins at room temperature. The detection reaction is stopped using 75ul of 1N NaOH. The egg white tests positive for significant levels 25 of the antibody.

25

**Example 25: Production of Interferon Using Avian Artificial Chromosomes in Quail**

Artificial chromosomes comprising a chicken (Barred-Rock) chromosome centromere (CAC) are prepared essentially using methods described in US Patent 30 No. 6,743,967, issued June 1, 2004, the disclosure of which is incorporated in its entirety herein by reference.

A coding sequence for interferon alpha 2b disclosed in US Patent Application No. 10/463,980, filed June 17, 2003, the disclosure of which is incorporated in its entirety herein by reference, is inserted in the 1OMC24-IRES1-Epo - ChromattB construct disclosed herein in Example 22 in place of the EPO coding sequence to produce 1OMC24-IRES-INF - ChromattB. The 1OMC24-IRES-INF - ChromattB is cloned into the CACs essentially as described in Example 23. The recombinant CACs are prepared then isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted to a concentration of about 10,000 chromosomes per  $\mu$ l of injection buffer. Approximately 50 nanoliters (nl) of injection buffer is injected per quail embryo.

Embryos for this study are collected from actively laying quail. Embryo donor quail are inseminated weekly using pooled semen from male quail of the same breed to produce eggs for injection.

On the day of egg collection, fertile quail are euthanized 2h post oviposition by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 40°C until microinjection.

Cytoplasmic injection of artificial chromosomes is achieved using the microinjection apparatus disclosed in US Patent Application No. 09/919,143, filed July 31, 2001. Chromosomes are injected into the Stage I embryos at a single site in each embryo.

Following microinjection, the embryos are transferred to the oviduct of recipient quail essentially as described in Olsen et al, B. J Exp Zool 109: 355-66, 1948. Typically, about one day after OT, the recipient quail lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until hatching of the birds.

Eggs from G2 transchromosomal quail test positive for the presence of intact functional interferon alpha 2b.

**Example 26: Production of Monoclonal Antibody Using Avian Artificial Chromosomes in Chicken**

An attB site and a hygromycin<sup>R</sup> coding sequence are inserted into the OMC24-IRES-LC and OMC24-IRES-HC vectors disclosed in US Patent Application No. 10/856,218, filed July 31, 2001, which are then each cloned into CACs of Example 25 essentially as described in Examples 22 and 23. The CACs are isolated by a dual laser high-speed flow cytometer.

The flow-sorted chromosomes are pelleted by centrifugation and are diluted to a concentration of 7000-12,000 chromosomes per  $\mu\text{l}$  of injection buffer. 10 Approximately 50 nanoliters (nl) of injection buffer is injected per chicken embryo.

Embryos for this study are collected from actively laying *G. gallus*. Embryo donor chickens are inseminated weekly using pooled semen from male chickens of the same breed to produce eggs for injection.

On the day of egg collection, fertile hens are euthanized 2h post oviposition 15 by cervical dislocation. The recently ovulated and fertilized eggs are collected from the upper magnum region of the oviduct under sterile conditions and placed in a glass well and covered with Ringers' Medium and maintained at about 41°C until microinjection.

Cytoplasmic injection of artificial chromosomes containing the OMC24- 20 IRES-LC is achieved using the microinjection apparatus disclosed US Patent Application No. 09/919,143. Approximately 500 chromosomes are injected into the Stage I embryos at a single site.

Following microinjection, the embryos are transferred to the oviduct of recipient chickens essentially as described in Olsen et al, B. J Exp Zool 109: 355- 25 66, 1948. Typically, about one day after OT, the recipient chickens lay a hard shell egg containing the manipulated ovum. Eggs are incubated in an incubator until hatching of the G0 birds.

G2 transchromosomal chickens are obtained which contain the artificial chromosome in their genome. The artificial chromosome containing the OMC24- 30 IRES-HC is introduced into embryos obtained from the G2 chickens in essentially the same manner as described for the OMC24-IRES-LC.

Eggs from G1 transchromosomal chickens which contain both the OMC-IRES-LC and OMC24-IRES-HC in their genome are tested for the presence of intact functional monoclonal antibody. A Costar flat 96-well plate is coated with 100 ul of C Goat-anti-Human kappa at a concentration of 5 µg/ml in PBS. The 5 plate is incubated at 37 °C for two hours. 200 µl of 5% PBA is added to the wells followed by an incubation at 37 °C for about 60-90 minutes followed by a wash. 100 ul of egg white samples (diluted in 1% PBA:LBP) is added to each well and the plate is incubated at 37 °C for about 60-90 min followed by a wash. 100 ul of a 1:2000 dilution of F'2 Goat anti-Human IgG Fc-AP in 1% PBA is added to the 10 wells and the plate is incubated at 37 °C for 60-90 min followed by a wash. The antibody is detected by placing 75 ul of 1mg/ml PNPP (p-nitrophenyl phosphate) in 5x developing buffer in each well and incubating for about 10-30 mins at room temperature. The detection reaction is stopped using 75ul of 1N NaOH. The egg white tests positive for significant levels of the antibody.

15

While this invention has been described with respect to various specific examples and embodiments, it is to be understood that the invention is not limited thereto and that it can be variously practiced with the scope of the following claims.

What is claimed is:

1. A transchromosomal avian.
- 5 2. The transchromosomal avian of claim 1 wherein the avian is a G1 transchromosomal avian.
3. The transchromosomal avian of claim 1 wherein the avian is a G2 transchromosomal avian.
- 10 4. The transchromosomal avian of claim 1 wherein the avian is a germline transchromosomal avian.
5. The transchromosomal avian of claim 1 wherein the avian is selected from the group consisting of chicken, quail and turkey.
- 15 6. The transchromosomal avian of claim 1 wherein the artificial chromosome comprises a centromere.
- 20 7. The transchromosomal avian of claim 6 wherein the centromere is an insect centromere.
8. The transchromosomal avian of claim 6 wherein the centromere is a mammalian centromere.
- 25 9. The transchromosomal avian of claim 6 wherein the centromere is an avian centromere.
- 30 10. The transchromosomal avian of claim 1 wherein the artificial chromosome comprises a heterologous nucleotide sequence.

11. The transchromosomal avian of claim 10 wherein the nucleotide sequence is heterologous to the avian.

12. The transchromosomal avian of claim 10 wherein the nucleotide  
5 sequence is heterologous to the chromosome.

13. The transchromosomal avian of claim 10 wherein the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance.

10 14. The transchromosomal avian of claim 10 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

15 15. The transchromosomal avian of claim 14 wherein the gene expression controlling region is a promoter which is operable in a cell of an oviduct.

20 16. The transchromosomal avian of claim 14 wherein the gene expression controlling region includes a promoter selected from the group consisting of a lysozyme promoter, an ovomucin promoter, an ovomucoid promoter and an ovalbumin promoter.

17. The transchromosomal avian of claim 1 wherein the avian lays an egg comprising a heterologous protein.

25 18. A method of producing a transchromosomal avian comprising:  
inserting a heterologous nucleotide sequence in a chromosome;  
substantially purifying the chromosome;  
introducing the chromosome into an avian embryo;  
maintaining the embryo under conditions suitable for the embryo to develop  
30 and hatch as a chick;  
thereby producing a transgenic avian.

19. The method of claim 18 wherein the avian can produce transchromosomal offspring.

20. The method of claim 18 wherein the chromosome is introduced by  
5 microinjection.

21. The method of claim 18 wherein the chromosome is introduced into the embryo by delivering the chromosome to a germinal disc.

10 22. The method of claim 18 wherein the embryo is an early stage embryo.

23. The method of claim 18 wherein the embryo is a stage I embryo.

15 24. The method of claim 18 wherein between 1 and about 10,000 chromosomes are introduced into the embryo.

25. The method of claim 18 wherein between 1 and about 1,000 chromosomes are introduced into the embryo.

20 26. The method of claim 18 wherein the avian is selected from the group consisting of chicken, quail and turkey.

27. The method of claim 18 wherein the chromosome comprises a  
25 heterologous nucleotide sequence.

28. The method of claim 27 wherein the heterologous nucleotide sequence includes a coding sequence for a therapeutic substance.

30 29. The method of claim 27 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

30. The method of claim 29 wherein the gene expression controlling region includes a promoter selected from the group consisting of a lysozyme promoter, an ovomucin promoter, an ovomucoid promoter and an ovalbumin promoter.

5

31. An avian cell comprising an artificial chromosome wherein the artificial chromosome includes a nucleotide sequence encoding a therapeutic substance.

10 32. The avian cell of claim 31 wherein the avian is selected from the group consisting of chicken, quail and turkey.

33. The avian cell of claim 31 wherein the heterologous nucleotide sequence includes a gene expression controlling region.

15

34. A method of producing a transgenic vertebrate animal comprising:

introducing into an embryo of a vertebrate animal a recombination site such that the recombination site is present in ovum of a mature vertebrate animal 20 developed from the embryo wherein the embryo does not normally comprise the recombination site,

introducing into the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal a nucleotide sequence comprising a second 25 recombination site and a coding sequence;

introducing into the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal a substance which facilitates insertion of the nucleotide sequence comprising a second recombination site and a coding sequence 30 proximal to the recombination site; and

exposing the embryo of the vertebrate animal or fertilized ovum of the mature vertebrate animal to conditions which lead to development of a viable transgenic vertebrate animal,

5 thereby producing a transgenic vertebrate animal.

35. The method of claim 34 wherein the vertebrate animal is an avian.

36. The method of claim 34 wherein the vertebrate animal is an avian  
10 selected from the group consisting of chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

37. The method of claim 34 wherein the vertebrate animal is a chicken.  
15

38. The method of claim 34 wherein the embryo or fertilized ovum is selected from the group consisting of a stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryo.

20 39. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by transfection.

25 40. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by microinjection.

30 41. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by cell fusion.

42. The method of claim 34 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by lipofection.

5 43. The method of claim 34 wherein the recombination site or nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum in the presence of PEI.

10 44. The method of claim 34 wherein at least one of the recombination site or the second recombination site is isolated from a bacteriophage.

45. The method of claim 34 wherein the recombination site is an attP site or an attB site.

15 46. The method of claim 34 wherein the second recombination site is an attP site or an attB site.

47. The method of claim 34 wherein the nucleotide sequence is stably incorporated into the genome of the embryo or fertilized ovum.

20 48. The method of claim 34 wherein the substance is an enzyme.

49. The method of claim 48 wherein the enzyme is a site specific recombinase.

25 50. The method of claim 48 wherein the enzyme is selected from the group consisting of a serine recombinase and a tyrosine recombinase.

30 51. The method of claim 48 wherein the enzyme is selected from the group consisting of EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1,  $\Phi$ 370.1,  $\Phi$ 105,  $\Phi$ FC1,

A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4,  $\Phi$ Rv1, Y4bA and Bja serine recombinases.

52. The method of claim 48 wherein the enzyme is selected from the group consisting of  $\Phi$ C31, TP901-1 and R4 serine recombinases.

53. The method of claim 48 wherein the enzyme is  $\Phi$ C31 serine recombinase.

10 54. The method of claim 34 wherein the substance is nucleic acid.

55. The method of claim 54 wherein the nucleic acid is DNA or RNA.

15 56. The method of claim 55 wherein a vector comprises the DNA.

57. The method of claim 55 wherein the nucleic acid encodes an enzyme.

20 58. The method of claim 55 wherein the nucleic acid encodes a site specific recombinase.

59. The method of claim 55 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine recombinase.

25 60. The method of claim 55 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1,  $\Phi$ 370.1,  $\Phi$ 105,  $\Phi$ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, 30 SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4,  $\Phi$ Rv1, Y4bA and Bja serine recombinases.

61. The method of claim 55 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of ΦC31, TP901-1 and R4 serine recombinases.

5 62. The method of claim 55 wherein the nucleic acid encodes a ΦC31 serine recombinase.

63. The method of claim 34 wherein the nucleotide sequence comprises an expression cassette.

10 64. The method of claim 34 wherein the coding sequence encodes a therapeutic composition.

15 65. The method of claim 64 wherein the therapeutic composition comprises at least one of a light chain or a heavy chain of an antibody.

66. The method of claim 64 wherein the therapeutic composition is a cytokine.

20 67. The method of claim 64 wherein the therapeutic composition is selected from the group consisting of interferon, erythropoietin, and granulocyte-colony stimulating factor.

25 68. The method of claim 34 wherein the coding sequence encodes a polypeptide present in an egg produced by the transgenic vertebrate animal wherein the animal is an avian.

69. An egg of claim 68.

30 70. The method of claim 34 wherein the recombination site is introduced into the embryo of a vertebrate animal or fertilized ovum before fertilization.

71. An ovum or sperm produced by a transgenic vertebrate animal of claim 34.

5 72. A transgenic vertebrate animal produced according to the method of claim 34.

73. A descendent of the transgenic vertebrate animal of claim 34.

10 74. A method of producing a transgenic vertebrate animal comprising:

introducing into an embryo of a vertebrate animal a recombination site such that the recombination site is present in sperm of a mature vertebrate animal developed from the embryo wherein the embryo does not normally comprise the recombination site,

15 fertilizing an ovum with sperm comprising the recombination site;

20 introducing into the ovum a nucleotide sequence comprising a second recombination site and a coding sequence and a substance which facilitates insertion of the nucleotide sequence comprising a second recombination site and coding sequence proximal to the recombination site; and

exposing the fertilized ovum to conditions which lead to development of a viable transgenic vertebrate animal,

25 thereby producing a transgenic vertebrate animal.

75. The method of claim 74 wherein the vertebrate animal is an avian.

30 76. The method of claim 74 wherein the vertebrate animal is an avian selected from the group consisting of chicken, turkey, duck, goose, quail,

pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary.

77. The method of claim 74 wherein the vertebrate animal is a chicken.

5

78. The method of claim 74 wherein the embryo is selected from the group consisting of a stage I, stage II, stage III, stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and stage XII embryo.

10 79. The method of claim 74 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by cell fusion.

15 80. The method of claim 74 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by lipofection.

20 81. The method of claim 74 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by transfection.

25 82. The method of claim 74 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the embryo or fertilized ovum by microinjection.

83. The method of claim 74 wherein the recombination site or nucleotide sequence comprising a second recombination site and a coding sequence is introduced in the presence of PEI.

30

84. The method of claim 74 wherein at least one of the recombination site or the second recombination site is isolated from a bacteriophage.

85. The method of claim 74 wherein the recombination site is an attP site or an attB site.

5 86. The method of claim 74 wherein the second recombination site is an attP site or an attB site.

87. The method of claim 74 wherein the nucleotide sequence is stably incorporated into the genome of the embryo or fertilized ovum.

10

88. The method of claim 74 wherein the substance is an enzyme.

89. The method of claim 88 wherein the enzyme is a site specific recombinase.

15

90. The method of claim 88 wherein the enzyme is selected from the group consisting of a serine recombinase and a tyrosine recombinase.

91. The method of claim 88 wherein the enzyme is selected from the  
20 group consisting of EcoYBCK, ΦC31, SCH10.38c, SCC88.14, SC8F4.15c,  
SCD12A.23, Bxb1, WwK, Sau CcrB, Bsu CisB, TP901-1, Φ370.1, Φ105, ΦFC1,  
A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24,  
SC2E1.37, SCD78.04c, R4, ΦRv1, Y4bA and Bja serine recombinases.

25 92. The method of claim 88 wherein the enzyme is selected from the group consisting of ΦC31, TP901-1 and R4 serine recombinases.

93. The method of claim 88 wherein the enzyme is ΦC31 serine recombinase.

30

94. The method of claim 74 wherein the substance is nucleic acid.

95. The method of claim 94 wherein the nucleic acid is DNA or RNA.

96. The method of claim 95 wherein a vector comprises the DNA.

5 97. The method of claim 94 wherein the nucleic acid encodes an enzyme.

98. The method of claim 94 wherein the nucleic acid encodes a site specific recombinase.

10 99. The method of claim 94 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine recombinase.

15 100. The method of claim 94 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK, ΦC31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1, Φ370.1, Φ105, ΦFC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4, ΦRv1, Y4bA and Bja serine recombinases.

20 101. The method of claim 94 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of ΦC31, TP901-1 and R4 serine recombinases.

25 102. The method of claim 94 wherein the nucleic acid encodes a ΦC31 serine recombinase.

30 103. The method of claim 74 wherein the nucleotide sequence comprises an expression cassette.

104. The method of claim 74 wherein the coding sequence encodes a therapeutic composition.

105. The method of claim 104 wherein the therapeutic composition  
5 comprises at least one of a light chain or a heavy chain of an antibody.

106. The method of claim 105 wherein the antibody is a human antibody.

107. The method of claim 104 wherein the therapeutic composition is a  
10 cytokine.

108. The method of claim 104 wherein the therapeutic composition is selected from the group consisting of interferon, erythropoietin, and granulocyte-colony stimulating factor.

15

109. The method of claim 74 wherein the coding sequence encodes a polypeptide present in an egg produced by the transgenic animal wherein the animal is an avian.

20

110. An egg of claim 109.

111. The method of claim 74 wherein the recombination site is introduced into the embryo of a vertebrate animal or fertilized ovum before fertilization.

25

112. An ovum or sperm produced by a transgenic vertebrate animal of claim 74.

30

113. A transgenic vertebrate animal produced according to the method of claim 74.

114. A descendent of the transgenic vertebrate animal of claim 74.

115. A method of modifying a vertebrate cell genome comprising:

introducing into a vertebrate cell a recombination site such that the  
5 recombination site is inserted into the vertebrate cell genome wherein the genome  
does not normally comprise the recombination site;

introducing a nucleotide sequence comprising a second recombination site  
and a coding sequence into the vertebrate cell or progeny cell thereof; and

10

introducing into the vertebrate cell or progeny cell thereof a substance  
which facilitates insertion of the nucleotide sequence comprising a second  
recombination site and a coding sequence proximal to the recombination site,

15 thereby modifying a vertebrate cell genome.

116. The method of claim 115 wherein the vertebrate cell is a cell of an  
avian.

20

117. The method of claim 115 wherein the vertebrate cell is a chicken  
cell.

118. The method of claim 115 wherein the vertebrate cell is a germ line  
cell.

25

119. The method of claim 115 wherein the vertebrate cell is an ovum or  
an embryo.

30

120. The method of claim 115 wherein the vertebrate cell an embryo cell  
and the embryo is selected from the group consisting of a stage I, stage II, stage III,  
stage IV, stage V, stage VI, stage VII, stage VIII, stage IX, stage X, stage XI and  
stage XII embryo.

121. The method of claim 115 wherein at least one of the recombination site or the nucleotide sequence comprising a second recombination site and a coding sequence is introduced into the vertebrate cell genome by a method selected  
5 from the group consisting of cell fusion, lipofection, transfection and microinjection.

122. The method of claim 115 wherein the recombination site or nucleotide sequence comprising a second recombination site and a coding sequence  
10 is introduced into the vertebrate cell genome in the presence of PEI.

123. The method of claim 115 wherein at least one of the recombination site or the second recombination site is isolated from a bacteriophage.

15 124. The method of claim 115 wherein the recombination site is an attP site or an attB site.

125. The method of claim 115 wherein the second recombination site is an attP site or an attB site.  
20

126. The method of claim 115 wherein the nucleotide sequence is stably incorporated into the genome of the vertebrate cell.

25 127. The method of claim 115 wherein the substance is an enzyme.

128. The method of claim 127 wherein the enzyme is a site specific recombinase.

30 129. The method of claim 127 wherein the enzyme is selected from the group consisting of a serine recombinase and a tyrosine recombinase.

130. The method of claim 127 wherein the enzyme is selected from the group consisting of EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1,  $\Phi$ 370.1,  $\Phi$ 105,  $\Phi$ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, 5 SC2E1.37, SCD78.04c, R4,  $\Phi$ Rv1, Y4bA and Bja serine recombinases.

131. The method of claim 127 wherein the enzyme is  $\Phi$ C31 serine recombinase.

10 132. The method of claim 115 wherein the substance is nucleic acid.

133. The method of claim 132 wherein the nucleic acid encodes an enzyme.

15 134. The method of claim 132 wherein the nucleic acid encodes a site specific recombinase.

135. The method of claim 132 wherein the nucleic acid encodes an enzyme selected from the group consisting of a serine recombinase and a tyrosine 20 recombinase.

136. The method of claim 127 wherein the nucleic acid encodes a serine recombinase selected from the group consisting of EcoYBCK,  $\Phi$ C31, SCH10.38c, SCC88.14, SC8F4.15c, SCD12A.23, Bxb1, WwK, Sau CcrB, BsU CisB, TP901-1, 25  $\Phi$ 370.1,  $\Phi$ 105,  $\Phi$ FC1, A118, Cac1956, Cac1951, Sau CcrA, Spn, TnpX, TndX, SPBc2, SC3C8.24, SC2E1.37, SCD78.04c, R4,  $\Phi$ Rv1, Y4bA and Bja serine recombinases.

137. The method of claim 127 wherein the nucleic acid encodes a  $\Phi$ C31 30 serine recombinase.

138. The method of claim 115 wherein the coding sequence encodes a therapeutic composition.

139. The method of claim 138 wherein the therapeutic composition  
5 comprises at least one of a light chain or a heavy chain of an antibody.

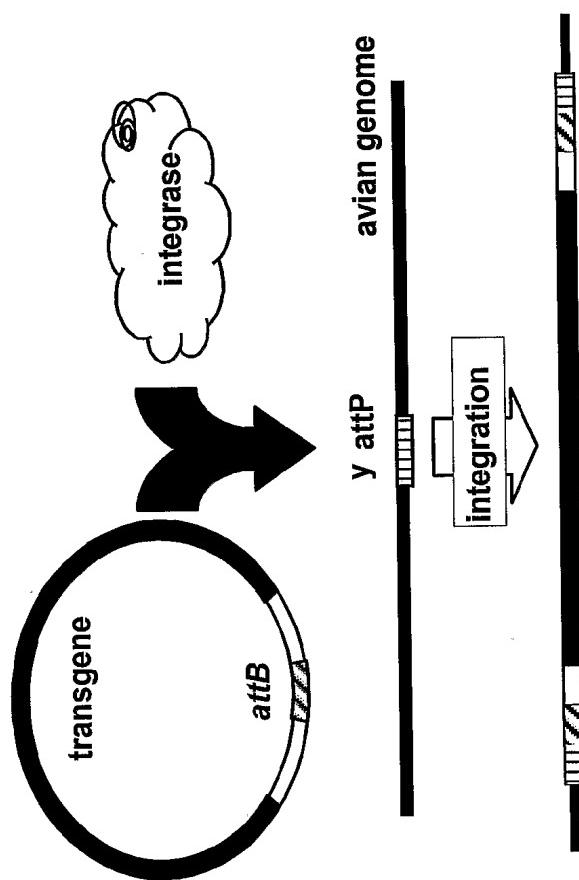
140. The method of claim 138 wherein the therapeutic composition is a cytokine.

10 141. The method of claim 138 wherein the therapeutic composition is selected from the group consisting of interferon, erythropoietin, and granulocyte-colony stimulating factor.

142. A cell produced according to the method of claim 115.

15

20



*Fig. 1*

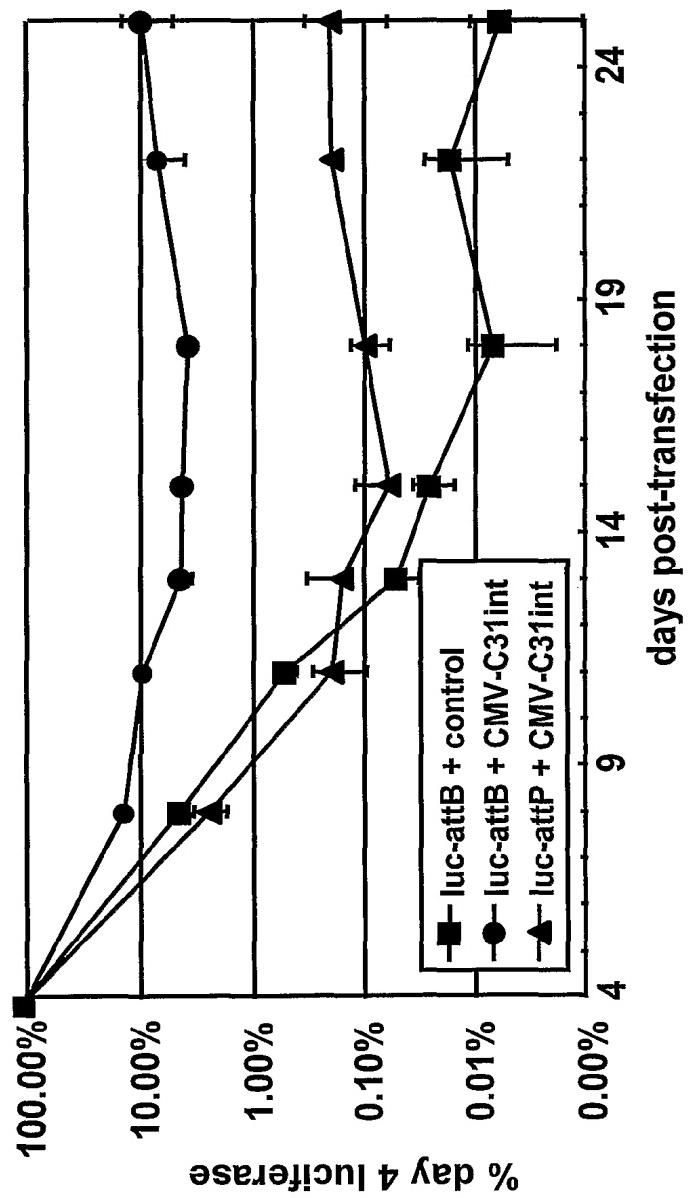
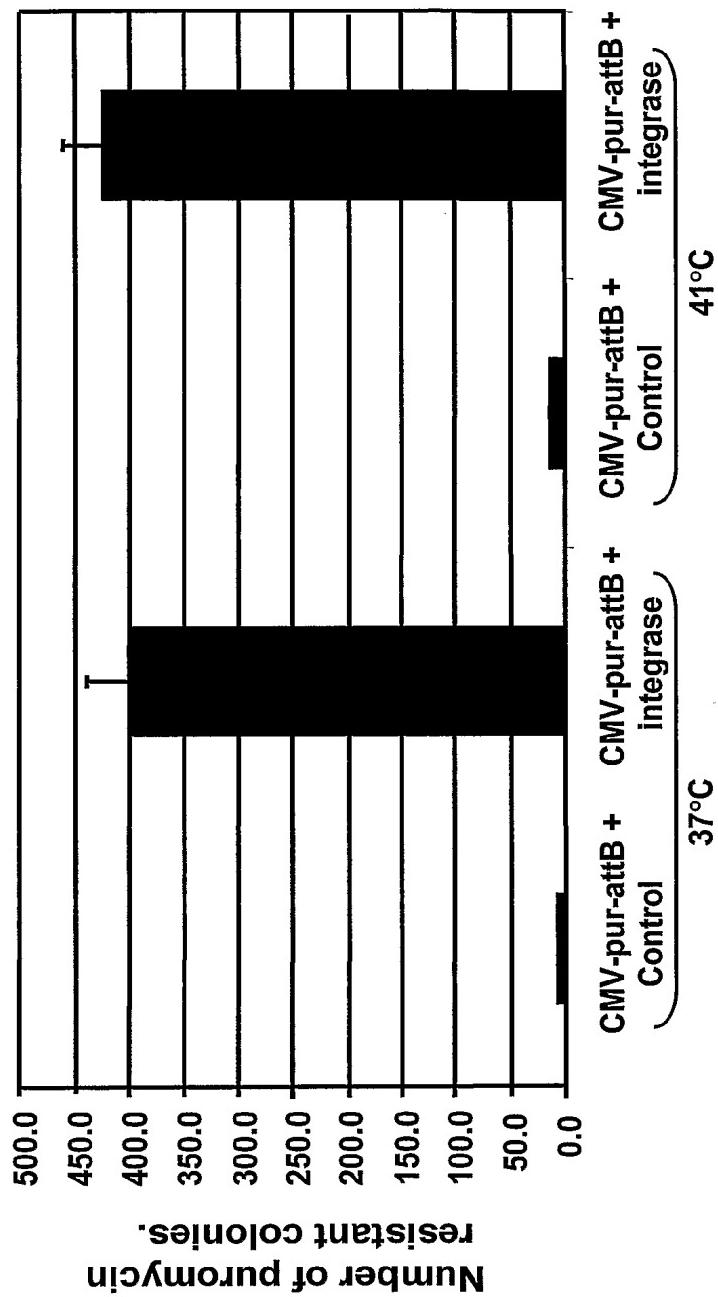
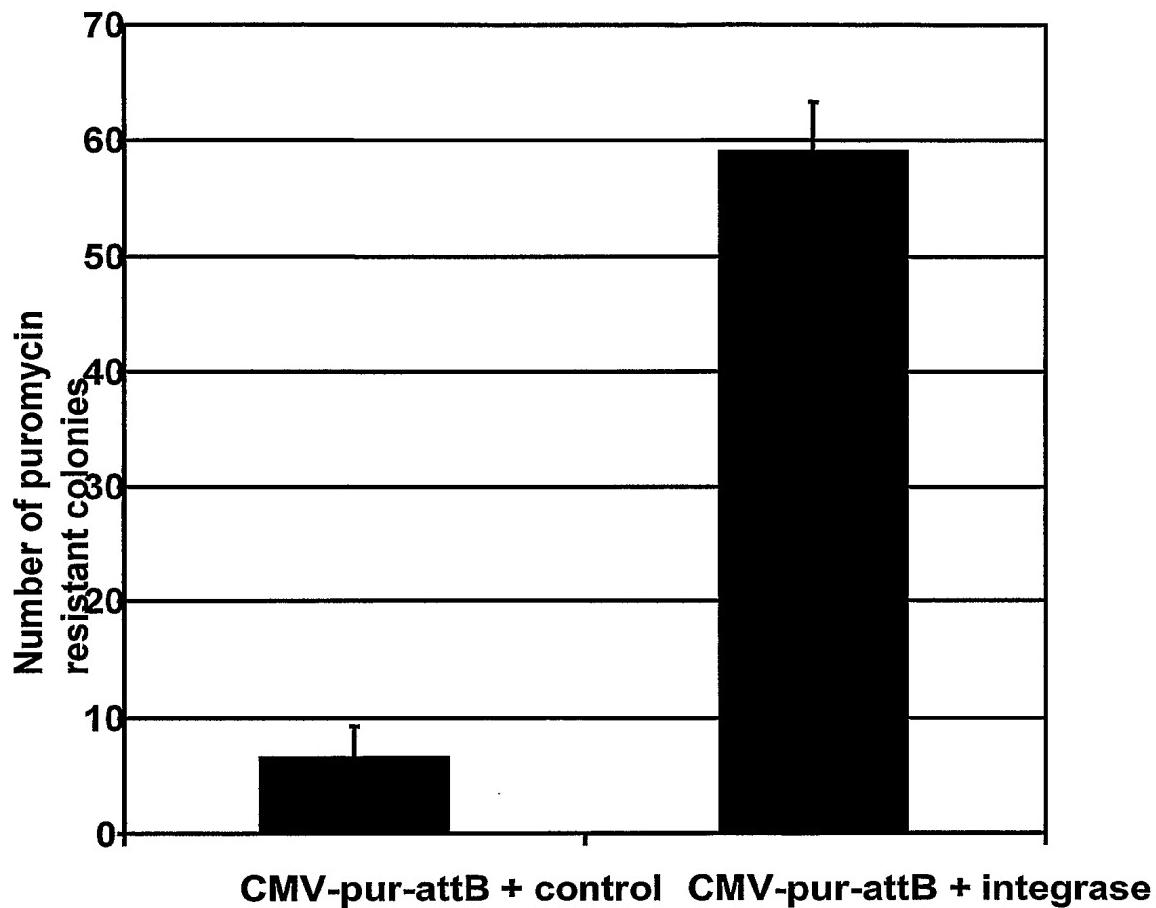


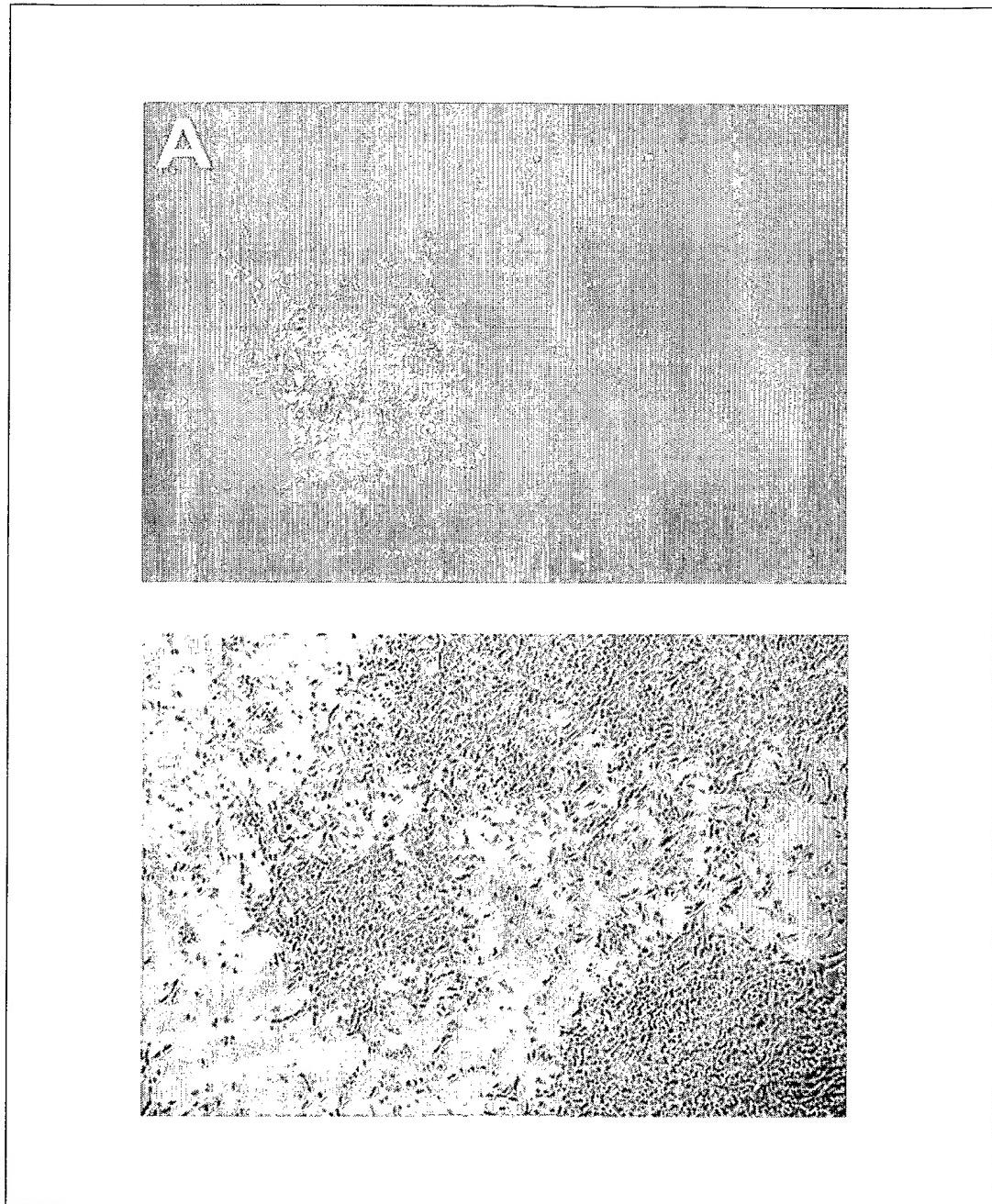
Fig. 2



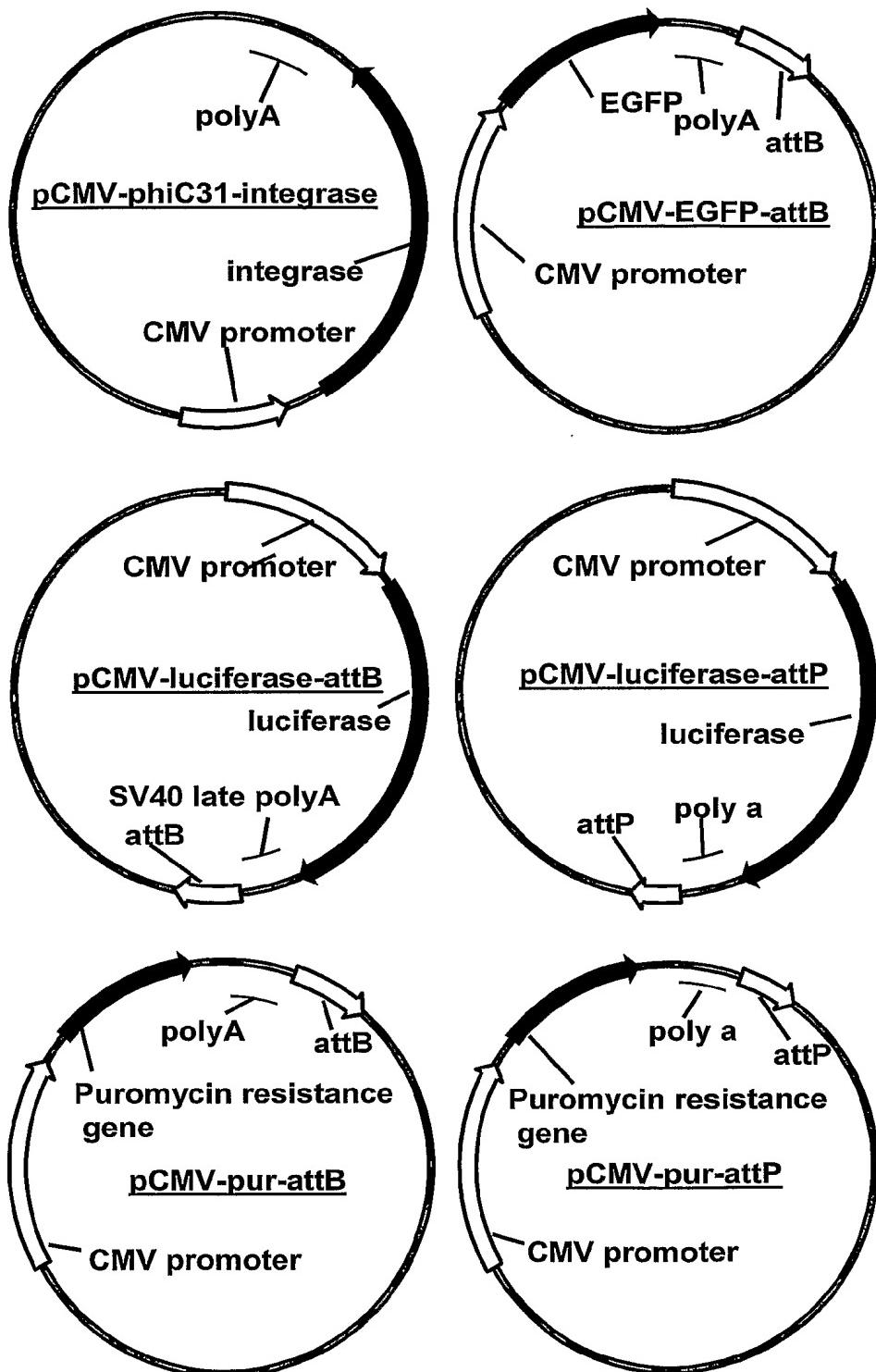
*Fig. 3*



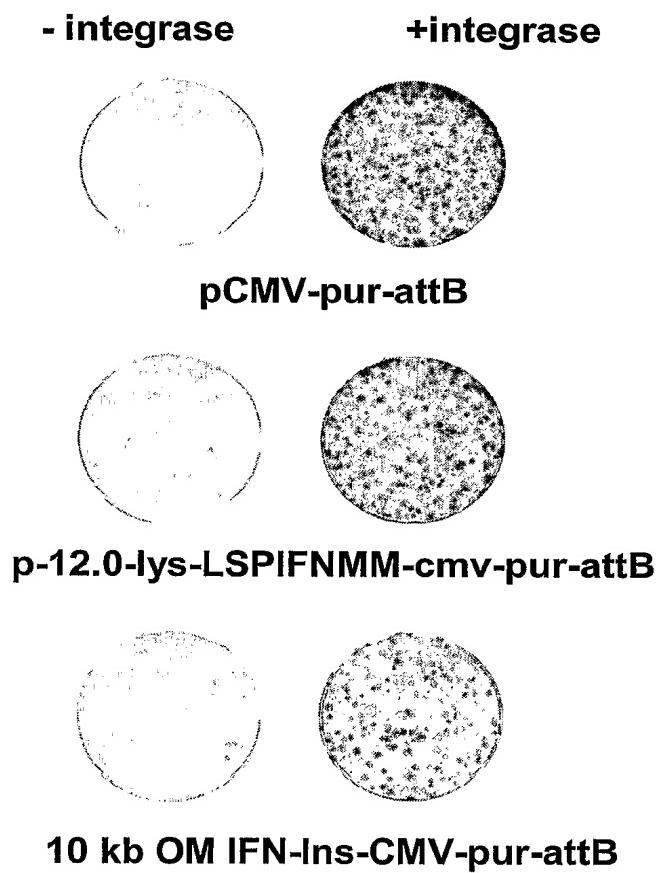
*Fig. 4*



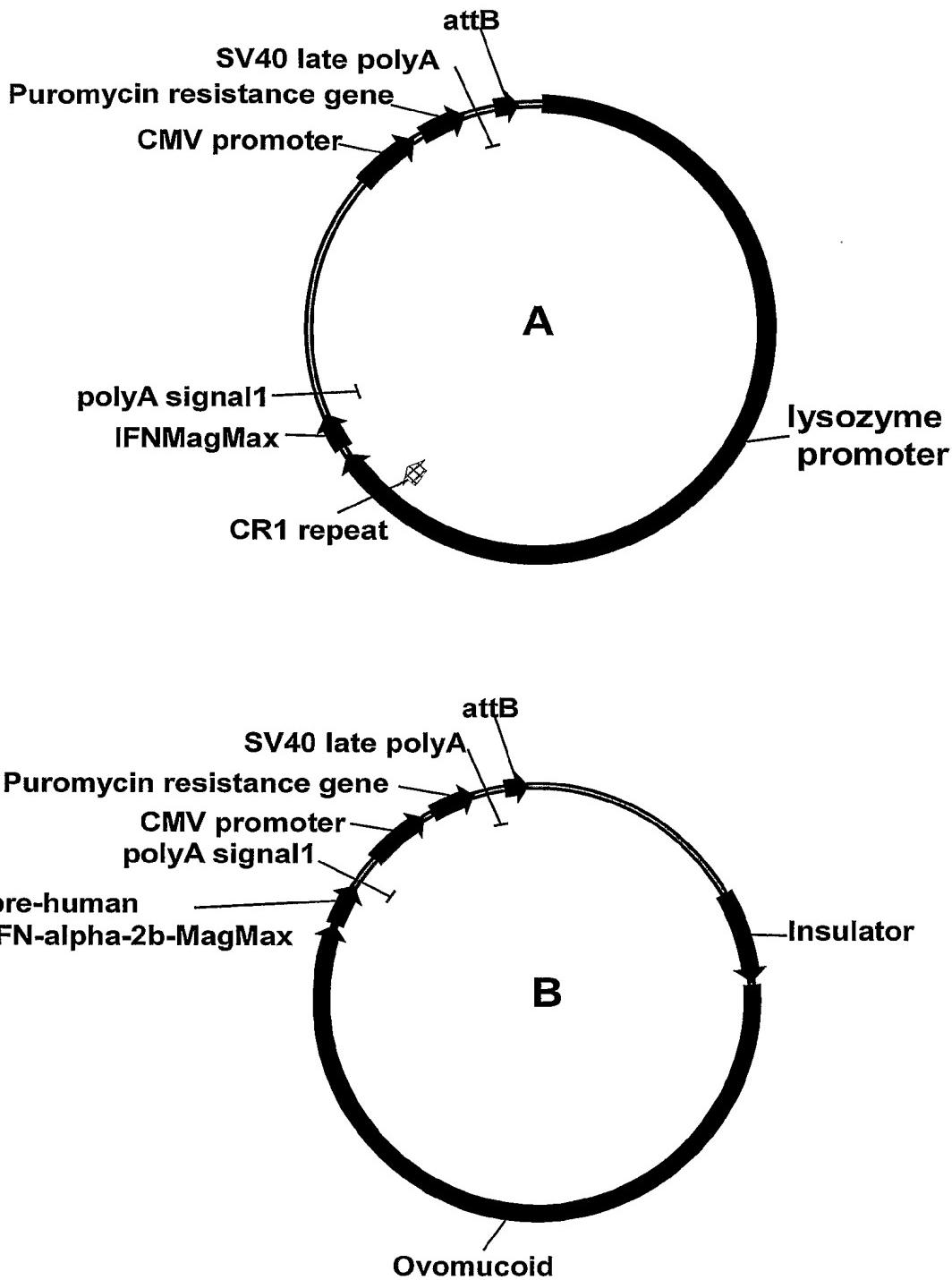
*Fig. 5*



*Fig. 6*



***Fig. 7***



**Fig. 8**

**pCMV-C31int (SEQ ID NO: 1)**

CATTGCCATT CAGGCT GCGCAACT GTGGGAAGGGC ATCGGT GCGGGC CTTCGCT ATT  
ACGCCAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAGCAGGATCG  
ATCCAGACATGATAAGATACTTGATGAGTTGGACAAACCACA ACTAGAATGCAGTGAAA  
AAATGCTTATTGTGAAATTGTGATGCTATTGCTTATTGTAAACCATTATAAGCTGCAA  
TAAACAAGTTAACAAACAATTGCATTCAATTATGTTCAAGGTTCAAGGGGAGGTGTGGG  
AGGTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCATGAACAG  
ACTGTGAGGACTGAGGGCCTGAAATGAGCCTTGGGACTGTGAATCTAAAATACACAAACAA  
TTAGAATCACTAGCTCCTGTGTATAATATTTCATAAAATCATACTCAGTAAGCAAACACTCTC  
AAGCAGCAAGCATATGCAGCTAGTTAACACATTATAACACTTAAAAATTATTTACCTT  
AGAGCTTAAATCTCTGTAGGTAGTTGTCATTATGTACACCCACAGAAGTAAGGTTCTT  
TCACAAAGATCCAAGCTAGCTTAAATACGACTCACTATAGGGAGAGAGCTATGACGTCG  
ATGCACCGTAAAGCTTGGGCCCCCTCGAGGGATCGGGGTGTCGCTACGCCGCTACGTCTC  
CGTGCCTCTCGATGGCGTCCCTGCCCCCTGCGCTAGTCGACTTCGTGACAACGATCTG  
TCTACGAAGAGGCCGACGAACACCGCGCTTGTGCGTCACTGACGCGCGCCCCCACGACTT  
AGGGCCGGTGGGTCAAGCGTCGGCGTCTTGGGAACCATGGTCAAGGGGAAGCTTGGGG  
CTTCGGCGCTTCAAGTTCGGCAAGCCGCTTCCGCCCCCTGCTGCGGAGCGTCAGCGCT  
GCCTGTTGCTTCGGAAAGTGCCTCCTGCCAACGGGTCGGTACGCGCCTGCGCGCG  
TTCGTACAGCTTCAAGGGCGTTCAAGGGCGTCTGCGCCTCCGCAACAAGGTTGCGCCGTT  
CGCGCTCTCTCAGGCCCTCAGTGAGCTTGCCTGCGAAGCGTCGGCGGCTTCCCACAGAAC  
GCCAACGTCTCTCGTCGCCCTCGCGTGCCTGATCTTGTGAAGATGCGTTCGCAACGAA  
CTTGTGAGTGCCTGCATGCTGACGTTGCACTGCGCTTGTGCTGCCCAGGTGCGGACGGT  
CGACCACTTCCGGCGACGGCAGCGTAAGAGTCCTGATCGATTCTCCCGCTTCGAA  
GTCATGACGGGCCACACTCGCAGTACAGCTTGTCCATGGCGGACAGAACGGCTTGGCG  
GGAAAGCCCCCTGCCGCCCTGCCGTCACCGCTGAAGCTCATACCAACTCAGCGG  
GCTCGATGATCGGTCCGCAATCAAGCTCGACCGGCCGGAGCGTGAATGGGGTCCGCTGAATG  
CGGTAACCTCAATCTCGTGGTGGCGTCCGGTCTCTCTGTAGATCACCTCAGC  
GGCGAAGCCCGCAATACGGGGTCCGAAGGATTCCGATAACGGTTGCCGGTCCCAGGCG  
TTGAAGCGGTCTTCTTCCAATCGTCTCGCCCCGGTCGGCACGGCGTCAGCGTCCATGCGC  
TTACAAAGCCCCGTGATGCTGCCGGTGAATGGGGCTTGACTGCCGGCTTGAAGGGAAAG  
GTGTTGTGCGTCTTGATCTCACGCCACCACCGGATTACGTGGCGTCAACTCGAAGG  
GTCCGTAAGGGAGTGGTCAAGTGGCAAGCTTGTGATGACGACATTGACCATTGGCG  
TTGCGCGTGAATCTCGTCTCGAAACAAAGCTCGAAGCCGTAAGGGCGCTTCCGCCGAC  
GTACCCGCCAATTGCGCTGAAGGTTCTCGTGTGAGAATCTCGCGACTTCAGCGAAG  
ATTCTTGTGCGACGCGTCAAGCCGATAATCAGGTGAATCAGGTCCATGACGTTCCCTGC  
CGGAAGACGCCCTCTGAGTGGAAACAATCGTCACGCCAGGGCGAGCAATTCCGAGACAAT  
CGGAATCGCGTCCATGACCTTCAGGCGCGAGAACCGCGACACGTATAGACAATGATCATGT  
TGAGCCGCCGGCGCGATTCGTCAGGATGCGTTGAACTCCGGCGCTCCGCCGCTCCG  
AACGCCACGTGCCCGCGCTTCGCTGAATGCCGACGAACCTGAACCGGCCCGTCGCG  
CTCGACTTCGCGCTGAAGGTGGCGCTTGTCTCGTGGCGCTACGCTGTGCGTGGG  
TTGCGCGTCAATTCTCGCGTCGCGCACTGACGGTGTGAAGCACCGCGTACGTGTC  
ACCCCGGTACAACCCCTTGTGTCATGTCGGCGACCCCTACGACTAGTGTGAGCTGTC  
GGAATTCCGGACCGGTACCTGCAAGGCGTACCTTCTATAGTGTACCTAAATAGCTTTGCA  
AAAGCTAGGCTAGAGTCCGGAGGCTGGATCGGTCCGGTGTCTCTATGGAGGTCAA  
GCGTGGATGGCGTCTCCAGGCGATCTGACGGTTCACTAAACGAGCTCTGCTTATAGACCT  
CCCACCGTACACGCCCTACCGCCATTGCGTCAATGGGGCGAGTTGTTACGACATTGGA  
AAGTCCCCTGATTTGGTGCACAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGG  
AAATCCCCGTGAGTCAAACCGCTATCCACGCCATTGATGTAAGTGCACAAACCGCATC  
TGGTAATAGCGATGACTAATACGTAGATGACTGCGCAAGTAGGAAAGTCCCAT  
TACTGGGCATAATGCCAGGCGGGCATTACCGTCATTGACGTCAATAGGGGGCGTACTTGG  
CATATGATACACTTGATGTAAGTGCACGGCAGTTACCGTAAATACTCCACCCATTGAC  
GTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTGACGT  
CGGGGGTGGCGGCTCAGGCCAGGCGGCCATTACCGTAAGTTATGTAACGACCTGCAC

GATGCTTCCGTGAAATTGTTATCCGCTCACATTCCACACATTATACGAGCCGAA  
GCTATAAAGTGTAAAGCCTGGGGCCTAATGAGTGAAGGGCCTCGTATACGCCTATTTT  
ATAGGTTAATGTCATGATAATAATGGTTCTTAGACGTAGGTGGCACTTTGGGGAAATG  
TGCAGCGAACCCCTATTGTTATTTCTAAATACATTCAAAATATGTATCCGCTCATGAGA  
CAATAACCCTGATAAAATGCTTCAATAATATTGAAAAACGCGGAATTGCAAGCTCTGCATT  
ATGAATCGGCCAACGCGGGGGAGAGGCCTTGCGTATTGGCGCTTCGCTTCCTCG  
TCACTGACTCGCTCGCTCGGTCGTCGGCGAGCGGTATCAGCTCACTCAAAGGGC  
GTAATAACGGTTATCCACAGAATCAGGGATAACGCAAGGAAAGAACATGTGAGCAGAAAAGGCA  
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTCCATAGGCTCCGCC  
CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA  
AGATACCAGGCCTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCC  
TACCGGATACCTGTCGCTTCTCCCTCGGAAGCGTGGCGCTTCTCAATGCTCACGCT  
GTAGGTATCTCAGTTGGTAGGTCGTCGCTCCAAGCTGGCTGTGACGAACCC  
GTTCAAGCCGACCGCTGCCTTATCCGTAACTATCGTCTTGAGTCCAACCCGTAAGACA  
CGACTTATGCCACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGG  
GTGCTACAGAGTCTTGAAGTGGGGCTAACTACGCTACACTAGAAGGACAGTATTGGT  
ATCTGCGCTCTGTAAGGCCAGTTACCTCGGAAAAAGAGTTGGTAGCTCTGATCCGGCAA  
ACAAACACCAGCTGGTAGCGGTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAA  
AAGGATCTAAGAAGATCTTGATCTTCTACGGGTCTGACGCTCAGTGGAACGAAAAC  
TCACGTTAAGGGATTGTCATGCCATAACTCGTATAGCATACATTATACGAAGTTATGG  
CATGAGATTCAAAAAGGATCTCACCTAGATCCTTAAATTAAATGAAGTTAAAT  
CAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA  
CCTATCTCAGCGATCTGCTATTCGTTATCCATAGTGCCTGACTCCCCGTCGTAGAT  
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCATGATAACCGCAGACCCAC  
GCTCACCGCTCAGATTATCAGCAATAAACCGCAGCGGAAGGGCCAGCGCAGAAGT  
GGTCCTGCAACTTATCCGCTCCAGCTATTAAATTGTTGCCGGAAAGCTAGAGTAAG  
TAGTTGCCAGTTAATAGTTGCGAACGTTGTTGCATTGCTACAGGCATCGTGGTGTAC  
GCTCGTGTGTTGGTATGGCTTATCAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGA  
TCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTCGGTCTCCGATCGTTGTAGAAGTAA  
GTTGCCAGTGTATCACTCATGGTTATGGCAGCAGTCATAATTCTCTTACTGTCATGC  
CATCCGTAAGATGCTTCTGTGACTGGTGAGTACTCAACCAAGTCATTGAGAATAGTGT  
ATGCCGCGACCGAGTTGCTCTGCCGGCGTCAATACGGGATAATACCGGCCACATAGCAG  
AACTTTAAAAGTGTCTCATATTGAAAACGTTCTCGGGCGAAAAGTCAAGGATCTTAC  
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGCAGCCACTGATCTCAGCATCTT  
ACTTCAACAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCCAAAAAGGGAAAT  
AAGGGCGACCGAAATGTTGAATACTCATACTCTCCCTTTCAATATTATTGAAGCATTT  
ATCAGGGTTATTGTCATGCCAGGGTGGGCACACATATTGATACCGCAGTCACCTACAC  
AGCACATAATTCAATGCGACTCCCTCTATCGCACATCTTAGACCTTATTCTCCCTCCAGC  
ACACATCGAAGCTGCCGAGCAAGCCGTTCTCACCAGTCCAAGACCTGGCATGAGCGGATA  
TATTGAAATGTTAGAAAATAACAAATAGGGTTCCGCGCACATTCCCCAAAAGTG  
CCACCTGAAATTGTAACGTTAATATTGTTAAAATTGCGTTAAATTGTTAAATCAG  
CTCATTGTTAACCAATAGGCCGAAATCGGAAAATCCCTATAAATCAAAGAACAGCC  
AGATAGGGTTGAGTGTGTTCCAGTTGGAACAAGAGTCCACTATTAAAGAACGTTGACTCC  
AACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATACCC  
ATCAAGTTTTGGGTGAGGTGCCGTAAGCACTAAACGCAACCTAAAGGGAGCCCC  
GATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAACGCAA  
GGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCCGTAACCACACCC  
CGCGCTTAATGCCCGTACAGGGCGCGTC

**Fig. 9**

**pCMV-luc-attB (SEQ ID NO: 2)**

CTCTATCGATAGGTACCGAGCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACTATTGAA  
TCAATATTGGCAATTAGCCATATTAGTCATGGTTATATAGCATAAATCAATATTGGCTATT  
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTATATTGGCTCATGTCCAATA  
TGACGCCATGTGACATTGATTATGACTAGTTATAATAGTAATCAATTACGGTAAATGGCCGCTGGCT  
AGTTCATAGCCCATAATATGGAGTCCCGTACATAACTACGGTAAATGGCCGCTGGCT  
GACCGCCAAACGACCCCCCCCATTGACGTCAATAATGACGTATGGTCCCATAGTAACGCA  
ATAGGGACTTCATTGACGTCAATGGTGGAGTATTTACGGTAAACTGGCCACTTGGCAGT  
ACATCAAGTGTATCATATGCCAAGTCCGCCCCATTGACGTCAATGACGGTAAATGGCCG  
CCTGGCATTATGCCAAGTACATGACCTTACGGACTTCCCTACTTGGCAGTACATCTACGTA  
TTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCG  
GTTGACTCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTGTTGGC  
ACAAAATCAACGGACTTCCAAAATGTCGAACAACTCCGCCCCATTGACGCAAATGGC  
GGTAGGCGTGTACGGTGGAGGTCTATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGC  
CTGGAGACGCCATCCACGCTGTTTGACCTCCATAGAACAGACCCGGGACCGATCCAGCCTCC  
CCTCGAAGCTCGACTCTAGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC  
TGTGGTAAAGCACCAGGAAGACGCCAAAACATAAAAGAACAGGCCCCGGCCATTCTATC  
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGACATAAGGCTATGAAGAGATACGCCCTGGTT  
CCTGGAACAATTGCTTTACAGATGCACATATCAGGGTGGACATCACTACGCTGAGTACTT  
CGAAATGTCGTTGGCAGAAGCTATGAAACGATATGGCTGAATAACAAATCACAGAA  
TCGTCGTATGCACTCTCTTCATTCTTATGCCGGTGTGGCGCGTATTATC  
GGAGTTGCAGTTGCGCCCGGAACGACATTATAATGAAACGTAATTGCTCAACAGTATGG  
CATTTCGAGCCTACCGTGGTCTCGTTCAAAAAGGGGTTGCAAAAAATTGAAACGTC  
AAAAAAAGCTCCAATCATCCAAAAAATTATTATCATGGATTCTAAACGGATTACCAAGGAA  
TTTCAGTCGATGTACACGTTCGTACATCTCATCTACCTCCGGTTTAAATGAATACGATT  
TGTGCCAGAGTCCTCGATAGGGACAAGACAATTGCACTGATCATGAACCTCTGGATCTA  
CTGGTCTGCCTAAAGGTGCGCTCTGCCTCATAGAACCTGCCTGCGTGAGATTCTCGCATGCC  
AGAGATCCTATTGGCAATCAAATCATTCCGGATACTGCGATTAAAGTGTGTTCCATT  
CCATCACGGTTTGGATGTTACTACACTCGGATATTGATATGTGGATTTCGAGTCGTCT  
TAATGTATAGATTGAAGAAGAGCTGTTCTGAGGAGCCTCAGGATTACAAGATTCAAAGT  
GCGCTGCTGGTCCAACCTATTCTCCTTCTCGCCAAAAGCACTCTGATTGACAATACGA  
TTTATCTAATTACACGAAATTGCTTCTGGTGGCGCTCCCTCTCTAAGGAAGTCGGGAAG  
CGGTTGCCAAGAGGTTCCATCTGCAGGTACAGGCAAGGATATGGCTCACTGAGACTACA  
TCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGTGGTAAAGTTCC  
ATTTTTGAAGCGAAGGTGTGGATCTGGATACCGGAAAACGCTGGCGTTAATCAAAGAG  
GCGAACCTGTGTGAGAGGTCCATTGATTATGTCGGTTATGAAACAATCCGGAAGCGACC  
AACGCCCTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTACTGGACGAAGA  
CGAACACTCTCATCGTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG  
CTCCCGCTGAATTGGAATCCATCTGCTCCAACACCCCAACATCTCGACGCAGGTGCGCA  
GGTCTTCCCGACGATGACGCCGGTGAACCTCCGCCCGTGTGTTGGAGCACGGAAA  
GACGATGACGGAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCGAAAAGT  
TGCAGGAGGAGTTGTGTTGTGGACGAAGTACCGAAAGGTCTTACCGAAAACCTCGACGCA  
AGAAAAATCAGAGAGATCCTCATAAGGCCAAGAAGGGCGGAAAGATCGCCGTGTAATTCTA  
GAGTCGGGGCGGGCGCCGGCTTCGAGCAGACATGATAAGATACTTGTGAGTTGGACAAA  
CCACAACATAGAATGCACTGAAATTGCTTATTGTGAAATTGTGATGCTATTGCTTAA  
TTTGTAAACCATTATAAGCTGCAATAAACAGTTAACACAACAAATTGCAATTGTT  
TCAGGTTCAAGGGGAGGTGTGGAGGTTAAAGCAAGTAAACCTCTACAAATGTGGTA  
AAATCGATAAGGATCAATTGGCTTCAAGGTACCGTCCACGATGAGGTACGGTCTCGAAGC  
CGCGGTGCGGGTGCAGGGCGTGCCTTGGCTCCCGGGCGCGTACTCCACCTCACCCATC  
TGGTCCATCATGATGAACGGGTGAGGTGGCGGTAGTTGATCCGGCGAACGCGCGGCG  
CGGGAAAGCCCTGCCCTCGAAACCGCTGGCGCGTGGTACGGGTGAGCACGGGACGTGCGA  
CGCGTCCGGCGGGTGCAGGATACGCGGGGCGCGTCAAGGGTTCTGACGGTACGGCG  
ATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTC  
CTTCCGGTGGCGGGCGGGCATGACTATCGTCGCCGACTTATGACTGTTCTTATCATGC

AACTCGTAGGACAGGTGCCGGCAGCGCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGTCGGCGTCCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTCCATAGGCTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTCCCTCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCGCTTACCGGATAACCTGTCCGCCCTTCTCCCTCGGAAGCGTGGCGTTCTCAATGCTCACGCTGTAGGTATCTCAGTTGGTGTAGGTCTCGTCAAGCTGGCTGTGACGAACCCCCCTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTATCGCCACTGGCAACGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTGAACTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTGATCCGGAAACAAACCAACCGCTGGTAGCGGTGGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCTTTGATCTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACACGTTAAGGGATTGGTTCATGAGATTATAAAAGGATCTCACCTAGATCCTTTAAATTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTGGTCTGACAGTTACAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGCTATTGTTCATCCATAGTGCTGACTCCCCGTCTGTAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAACCGCGAGACCCA CGCTCACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGGAAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCATCCAGTCTATTAAATTGGTCCGGGGAAAGCTAGAGTAA GTAGTTGCCAGTTAATAGTTGCCAACGTTGTCCTGCACTACAGGCATCGTGGTGTCA CGCTCGCGTGGTATGGCTTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTCGGTCCGATCGTTGTCAAGAGTAAAGTTGGCGCAGTGTATCACTCATGGTTATGGCAGCACTGCAATAATTCTTACTGTCTGCCATCGTAAGATGCTTTCTGTGACTGGTAGACTCAACCAAGTCATTCTGAGAATAGTG TATGCGCGACCGAGTTGCTCTTCCCAGGCTCAATACGGGATAATACCGCGCCACATAGCAAACTTAAAGTGTCTCATATTGAAACGTTCTCGGGGGAAACACTCTCAAGGATCTTACCGCTGAGATCCAGTTCGATGTAACCCACTCGTGACCCAACTGATCTCAGCATCTTTACTTTCAACAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAAAGGGATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATTGAAGCATT TATCAGGGTTATTGTCTCATGAGCGGATACATATTGAAATGTATTAGAAAAAATAAACAAATAGGGGTCCCGCACATTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGATTAA GCGCGGGGGGTGGTGGTACGCGCAGCGTACACCTGCCAGCGCCCTAGCGCCCCGCTCTTCCTCGCTTCTTCCCTTCTCGCCACGTTCGCCGGCTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTCTGGTCCGATTAGGTTAGTGTCTTACGGCACCTGACCCCAAAAC TTGATTAGGGTGATGGTTCACGTAGTGGGCATGCCCTGATAGACGGTTTCGCCCCCTGTAGCGGCGATTAAACGTTGGACTCTTGTGAAACTGGACAAACACTCAACCC TATCTCGGTCTATTCTTGATTTATAAGGGATTGGCGATTTCGCGCTATTGTTCAAGCTTAACTTAAACAAATATTAAACGTTACAATTCCATTCCGCCATTCAAGGCTGCGCACATTCAAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTCGCTATTACGCCAGCCCAAGCTACCATGATAAGTAATATTAAAGGTACGGGAGGTACTGGAGCGGCGCAATAAAATATCTTATTTCATTACATCTGTGTGTTGGTTGTGAATCGATAG TACTAACATACGCTCTCCATCAAAACAAACGAAACAAACAAACTAGCAAAATAGGCTGTC CCCAGTGCAAGTGCAGGTGCCAGAACATT

**Fig. 10**

**pCMV-luc-attP (SEQ ID NO: 3)**

CTCTATCGATAGGTACCGAGCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAA  
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT  
GCCATTGCATACGGTATCTATATCATAAATATGTACATTATATTGGCTCATGTCCAATA  
TGACCGCCATGTGACATTGATTATGACTAGTTATAATAGTAATCAATTACGGTAAATGGCCCTGGCT  
AGTTCATAGCCATATATGGAGTCCGCGTACATAACTACGGTAAATGGCCCTGGCT  
GACCGCCAACGACCCCCGCCATTGACGTCATAATGACGTATGTTCCCATAGTAACGCCA  
ATAGGGACTTCATTGACGTCATGGGTGGAGTATTACGGTAAACTGCCACTGGCAGT  
ACATCAAGTGTATCATATGCCAAGTCCGCCCCATTGACGTCATGACGGTAAATGGCCCG  
CCTGGCATTATGCCAGTACATGACCTAACGGACTTCCCTACTTGGCAGTACATCTACGTA  
TTAGTCATCGCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCG  
GTTGACTCACGGGATTCCAAGTCTCCACCCCCATTGACGTCATGGGAGTTGTTGGC  
ACAAAATCAACGGACTTCAAAATGCGTAACAACCTCCGCCCCATTGACGCAAATGGG  
GGTAGGCGTGTACGGGGAGGTCTATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGC  
CTGGAGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCC  
CCTCGAAGCTCGACTCTAGGGCTCGAGATCTGCGATCTAAGTAAGCTGGCATCCGGTAC  
TGTTGGTAAAGCACCAGGAAAGACGCCAAAACATAAAAGAAAGGCCGGCGCATTCTATC  
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATAACGCCCTGGTT  
CCTGGAACAATTGCTTTACAGATGCACATATCGAGGGTGGACATCACTACGCTGAGTACCT  
CGAAATGTCGCTCGGTTGGCAGAAGCTATGAAACGATATGGCTGAATACAAATCACAGAA  
TCGTCGTATGCACTGAAAACCTCTCTCAATTCTTATGCCGGTTGGCGCTTATTATC  
GGAGTTGCAGTTGCGCCCGGAACGACATTATAATGAAACGTAATTGCTCAACAGTATGGG  
CATTTCGCAACCTTACCGTGGTGTTCGTTCCAAAAAGGGGTTGCAAAAAATTGAAACGTGC  
AAAAAAAGCTCCAATCATCCAAAAAAATTATTATCATGGATTCTAAACGGATTACCAAGGGA  
TTTCAGTCGATGTACACGTTCGTCACATCTCATCTACCTCCGGTTTAAATGAATACGATT  
TGTGCCAGAGTCCTCGATAGGGACAAGACAATTGCACTGATCATGAAACTCCTCTGGATCTA  
CTGGCTGCTAAAGGTGTCGCTCGCCTCATAGAAACTGCCTGCGTGAGATTCTCGCATGCC  
AGAGATCCTATTGGCAATCAAATCATTCCGGATACTGCGATTAAAGTGTGTTCCATT  
CCATCACGGTTTGGATGTTACTACACTCGGATAATTGATATGTGGATTTCGAGTCGCT  
TAATGTATAGATTGAAAGAAGAGCTGTTCTGAGGGCCTTCAGGATTACAAGATTCAAAGT  
GCGCTGCTGGTCCAACCTTATTCTCCTCTCGCCAAAAGCACTCTGATTGACAATACGA  
TTTATCTAATTACAGAAATTGCTCTGGTGGCGCTCCCTCTCTAAGGAAGTCGGGGAAG  
CGGTTGCAAGAGGTTCCATCTGCAAGGATCAGGCAAGGATATGGCTCACTGAGACTACA  
TCAGCTATTCTGATTACACCCGAGGGGATGATAAAACGGCGCGTGGTAAAGTTCC  
ATTTTTGAAGCGAAGGGTGTGGATCTGGATACCGGAAACGCTGGCGTTAATCAAAGAG  
GCGAACTGTGTGAGAGGTCTCATGATTATGTCGGTTATGTAACAAACATCCGGAAGCGACC  
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTACTGGGACGAAGA  
CGAACACTCTCATCGTGACCGCTGAAGTCTCTGATTAAAGTACAACAGGCTACAGGTGG  
CTCCCGCTGAATTGGAATCCATCTGCTCCAAACACCCAAACATCTCGACGCAGGTGTCGCA  
GGTCTCCCGACGATGACGCCGGTGAACCTCCGCCGTTGTTGGAGCACGGAAA  
GACGATGACGAAAAGAGATCGTGGATTACGTCGCCAGTCAGTAACAAACCGCGAAAAAGT  
TGC CGGAGGGAGTTGTGTTGGACGAAGTACCGAAAGGTCTACCGAAAACCTCGACGCCA  
AGAAAAATCAGAGAGATCCTCATAAAGCCAAGAAGGGCGGAAAGATGCCGTGTAATTCTA  
GAGTCGGGGCGGCCGGCCGTTCGAGCAGACATGATAAGATAACATTGATGAGTTGGACAAA  
CCACAACATAGAATGCACTGAAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCTTAA  
TTTGTAAACCATTATAAGCTGCAATAAACAGTTAACACAACAATTGCAATTCTATTGTT  
TCAGGTTCAAGGGGAGGTGTGGAGGTTTTAAAGCAAGTAAACCTCTACAAATGTGGTA  
AAATCGATAAGGATCAATTGCGCTCGACTAGTACTGACGGACACACCGAAGCCCCGGCGGC  
AACCCCTAGCGGATGCCCGGGCTTCACGTTTCCCAGGTAGAAGCGGTTTCCGGAGTA  
GTGCCCAACTGGGTAACCTTGTAGTTCTCAGTGGGGCGTAGGGTCCGACATGAC  
ACAAGGGTTGTGACCGGGTGGACACGTACGCGGGTGTACGACCGTCAGTCGCGCAGC  
GCGACTAGTACAAGCGAATTGATCCGTCACCGATGCCCTGAGAGCCTCAACCCAGTC  
GCTCCTCCGGTGGCGGGCATGACTATCGTCGCCACTTATGACTGTCTCTTATC  
ATGCAACTCGTAGGACAGGTGCCGAGCGCTTCCGCTTCGCTACTGACTCGCTGC

GCTCGGTCGTTGGCTGCCGGAGCGGTATCAGCTCACTCAAAGGCCTAATACGGTTATCC  
ACAGAACATCAGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA  
CCGTAAAAAGGCCGCGTGTGGCTTCCATAGGCTCCGCCCTGACGAGCATCAC  
AAAATCGACGCTCAAGTCAGAGGTGGCGAACCCGACAGGACTATAAAGATACCAGGCGTT  
CCCCCTGGAAGCTCCCTCGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTC  
CGCCTTCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTGACGAACCCCCGTTCAGCCCACCGC  
CGGTGTAGGTGCGTCCAAGCTGGGCTGTGACGAACCCCCGTTCAGCCCACCGC  
TGCGCCTATCCGTAACATCGTCTGAGTCCAACCCGTAAGACACGACTTATGCCACT  
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGGTATGTAGGCGGTACAGAGTTCT  
TGAAGTGGTGGCTTAACACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTG  
AAGCCAGTTACCTCGGAAAAGAGTTGGTAGCTCTGATCCGGCAAACAAACCACCGCTGG  
TAGCGGTGGTTTTGTTGCAAGCAGATTACCGCAGAAAAAAAGGATCTCAAGAAG  
ATCCTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATT  
TTGGTCATGAGATTATCAAAAGGATCTCACCTAGATCCTTTAAATTAAAAATGAAGTT  
TAAATCAATCTAAAGTATATGAGTAAACTTGGCTGACAGTTACCAATGCTTAATCAGTG  
AGGCACCTATCTCAGCGATCTGTCTTCCATAGTTGCCTGACTCCCCGTCGTG  
TAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA  
CCACGCTCACCGGCTCCAGATTATCAGCAATAAACCGAGCCAGCCGGAAAGGGCCAGCGCA  
GAAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGGCCGGAAAGCTAGA  
GTAAGTAGTCGCCAGTTAATAGTTGCGAACGTTGCTACAGGCTACGGCATCGTGGT  
GTCACGCTCGTGTGTTGGTATGGCTTCAATTGAGCTCCGGTCCAAAGGATCAAGGCGAGTTA  
CATGATCCCCATGTTGCAAAAAGCGGTTAGCTCCTCGGTCTCGATCGTTGTCAGA  
AGTAAGTTGGCCGAGTGTATCAGTGGTTATGCGAGCAGTGCATAATTCTCTACTGT  
CATGCCATCGTAAGATGCTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATTGAGAAT  
AGTGTATGCCGGGACCGAGTTGCTTGGCCGGCTCAATACGGGATAATACCGGCCACAT  
AGCAGAACTTAAAGGTCATCATTGAAAACGTTCTCGGGCGAAAACCTCAAGGAT  
CTTACCGCTGTTGAGATCCAGTCCAGTGTAAACCACTCGTGCACCCAACTGATCTTCAGCAT  
CTTTACTTACCGAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAATGCCGAAAAAG  
GGAATAAGGGCGACCGGAAATGTTGAATACTCATCTTCCCTTTCAATATTATTGAAG  
CATTTATCAGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAAAC  
AAATAGGGTTCCCGCAGCATTCCCGAAAAGTGCACCTGACGCGCCCTGTAGCGCGCA  
TTAAGCGCGGGGTGTGGTTACGCGCAGCGTACACTGCCCAGCTCGCCGGCTTCCCGTCAAG  
GCCGCTCCTTCGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCGTCAAG  
CTCTAAATCGGGGCTCCCTTCTGGGTTCCGATTAGTGTCTTACGGCACCTCGACCCCAA  
AAACTGATAGGGTGTGGTTACGTTCTTAATAGTGGACTCTGTTCAAACGGTAAACACTCA  
ACCCATCTCGGTCTATTCTTTGATTTAAAGGATTGCGGATTTGCCCTATTGGTTA  
AAAAATGAGCTGATTAACAAAATTAAACGCAATTAAACAAATATTACGTTACAAT  
TTCCCATCGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCTTCCG  
TATTACGCCAGCCAAGCTACCATGATAAGTAAGTAATATTAAAGGTACGGGAGGTACTTGG  
GCGGCCGAATAAAATATCTTATTTCATTACATCTGTGTGGTTTTGTGTGAATCG  
ATAGTACTAACATACGCTCTCCATCAAACAAAAGAAACAAACAAACTAGCAAAATAGGC  
TGTCCCCAGTGCAAGTGCAGGTGCCAGAACATT

*Fig. 11*

**pCMV-pur-attB (SEQ ID NO: 4)**

CTAGAGTCGGGGCGGCCGCGCTCGAGCAGACATGATAAGATAACATTGATGAGTTGGAC  
AAACCACAACAGAACATGCAAGTAAAAAATGCTTATTGTGAAATTGTGATGCTATTGCT  
TTATTTGTAACCATTATAAGCTGCAATAACAAGTTAACAAACAATTGCATTCTATTAT  
GTTTCAGGTTCAGGGGAGGTGTGGGAGGTTAAAGCAAGTAAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTCCGCTTCAGGTACCGTCAGGATGTAGGTACGGTCTCGA  
AGCCGGGTGCGGGTGCCAGGGCGCCCTGGGCTCCCCGGCGCGTACTCCACCTCACCC  
ATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGCGAACGCGCG  
CACCGGGAAAGCCCTGCCCTGAAACCGCTGGCGCGGTGGTACGGTGAGCACGGACGTG  
CGACGGCGTCGGCGGTGCGGATACGGGGCAGCGTCAGCGGTTCTGACGGTACGGCG  
GGCATGTCGACAGCGAATTGATCGTCGACCGATGCCCTGAGAGCCTCAACCCAGTCAG  
CTCCTCCGGTGGGCGCGGGCATGACTATCGTCGCCGCACTTATGACTGTTCTTATCA  
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTTCCGCTTCGCTCACTGACTCGCTGCG  
CTCGGTCGTTCGCTGCGCGAGCGGTATCAGCTCAAAAGGCGGTAAACGTTATCCA  
CAGAATCAGGGATAACGCAAGGAAACATGTGAGCAGAAAAGGCCAGCAAAAGGCCAGGAAC  
CGTAAAAAGGCCGCTTGTGGCGTTTCCATAGGCTCCGCCCTGACGAGCATCACAA  
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCGAGCGTTTC  
CCCCCTGGAAAGCTCCCTCGTGCCTCTCTGACCCCTGCCGTTACCGGATAACCTGTCC  
GCCCTTCTCCCTCGGGAAAGCGTGGCGCTTCTCAATGTCACGCTGTAGGTATCTCAGTT  
GGTAGGTCGTTCGCTCCAAGCTGGCTGTGTCACGAACCCGGTTCAAGACACGACTTATCGCCACTG  
GCGCCTTATCGGTAACACTCGTCTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG  
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT  
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGCCTGCTGA  
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTGTGATCCGGCAAACAAACCCACCGCTGGT  
AGCGGTGGTTTTTTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTAAGAAGA  
TCCTTGATCTTTCTACGGGGCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT  
TGGTCATGAGATTATCAAAAGGATCTTCACTAGATCCTTTAAATTAAAAATGAAGTTT  
AAATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA  
GGCACCTATCTCAGCGATCTGCTATTGTTCATCCATAGTTGCTGACTCCCGTCGTGT  
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATAACCGCGAGAC  
CCACGCTCACCGGCTCCAGATTATCAGCAATAAACCGCCAGCCGAAGGGCCGAGCGCAG  
AAAGTGGCCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTCGGGAGCTAGAG  
TAAGTAGTTGCCAGTTAATAGTTGCGCAACGTTGTCATTGCTACAGGCATCGTGGTG  
TCACGCTCGTGTGGTATGGCTCATTGAGCTCCGGTCCAAACGATCAAGGGAGTTAC  
ATGATCCCCATGTTGTGAAAAAGCGGTTAGCTCCTCGGCTCCGATCGTTGTCAGAA  
GTAAGTGGCCGAGTGTATCACTCATGGTATGGCAGCACTGCAATAATTCTTACTGTC  
ATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCATTCTGAGAATA  
GTGTATCGGGCGACCGAGTTGCTCTGCCCGCGTCAATAACGGATAATACCGGCCACATA  
GCAGAACTTTAAAGTGTCACTATTGAAAACGTTCTCGGGCGAAAACCTCTCAAGGATC  
TTACCGCTGTTGAGATCCAGTGTGATGTAACCCACTCGTCACCCACTGATCTCAGCATC  
TTTACTTCAACAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAAAGG  
GAATAAGGGCGACCGAAATGTTGAATACTCATACTCTTCAATATTATTGAAGC  
ATTATCAGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAATAACA  
AATAGGGTTCCGCGCACATTCCCGAAAAGTGCACCTGACCGCCCTGTAGCGCGCAT  
TAAGCGGGCGGGTGTGGTTACGCGCAGCGTACCGCTACACTTGCACGCCCTAGCG  
CCCGCTCCTTCCGCTTTCTCCCTCCTTCTCGCCACGTTGCCGCGCTTCCCGTCAGC  
TCTAAATCGGGGCTCCCTTACGGGTCCGATTTAGTGCTTACGGCACCTGACCCAAAA  
AACTTGATTAGGGTGTGATGGTACCGTAGTGGCCATGCCCTGATAGACGGTTTCTGCCCT  
TTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTGTTCCAAACTGGAACACACTCAA  
CCCTATCTCGGTCTATTCTTTGATTATAAGGGATTGCGGATTGCCCTATTGGTTAA  
AAAATGAGCTGATTAAACAAAATTAACGGAATTAAACAAAATTAACGTTACAATT  
TCCCATTGCCATTAGGCTGCGCAACTGTGGGAAGGGCGATCGGTGCGGCCCTTCTCGCT  
ATTACGCCAGCCAAAGCTACCATGATAAGTAAGTAATTAAAGGTACGGGAGGTACTTGGAG  
CGGCCGCAATAAAATATCTTATTTCATTACATCTGTTGTTGGTTTGTGAAATCGA

TAGTACTAACATACGCTCTCCATCAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT  
GTCCCCAGTGCAGTGCAAGTCAGGTGCCAGAACATTCTATCGATAGGTACCGAGCTTACGC  
GTGCTAGCCCTCGAGCAGGATCTATACTACATTGAATCAATATTGGCAATTAGCCATATTAGTCA  
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCATTGCATACGTTGTATCTATATCAT  
AATATGTACATTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAAATAGTAATCAATTACGGGGTCAATTAGTCATAGCCCATAATGGAGTTCCGCG  
TTACATAACTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCCGCCATTGACG  
TCAATAATGACGTATGTTCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGT  
GGAGTATTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC  
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAAGTACATGACCTTA  
CGGGACTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGTGCG  
GTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGGATTCCAAGTCTCC  
ACCCCATGACGTCAATGGGAGTTGTTGGCACCAAAATCAACGGGACTTCCAAAATGT  
CGTAACAACCTCCGCCATTGACGCAAATGGCGGTAGGCAGTACGGTGGGAGGTCTATAT  
AAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTGACC  
TCCATAGAAGACACCAGGACCGATCCAGCCTCCCTCGAAGCTGACTCTAGGGCTCGAGA  
TCTCGATCTAAGCTTGCATGCCCTGCAGGTGGCCACGACCGGTGCCACC  
CCCCCTGACCCACGCCCTGACCCCTCACAAAGGAGACGACCTTCCATGACCGAGTACAAGGCC  
ACGGTGCCTGCCACCCCGCACGACGACGCTCCCCGGCGTAGCACCCTGCCGCCGCGTT  
CGCCGACTACCCCGCCACCGCACCGTCGACCCGGACCGCCACATCGAGCGGGTCAACCG  
AGCTGCAAGAACTCTTCCCTCACGCCGTGGGCTCGACATCGGCAAGGTGTGGGTCGCCG  
GACGGCGCCGCGGTGGCGGTCTGGGACCGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGC  
CGAGATCGGCCCGCGCATGGCGAGTTGAGCGGTTCCCGCTGGCCGCCAGCAACAGATGG  
AAGGCCCTCTGGCGCCGACCGGGCCAAGGGAGGCCGCTGGGTCCTGGCCACCGTCGCCG  
TCGCCCCGACCAACCAGGGCAAGGGTCTGGGAGCGCCGTGCTCCCCGGAGTGGAGGCC  
CGAGCGCGCCGGGTGCCGCTTCTGGAGACCTCGCGCCCGCAACCTCCCCCTCTACG  
AGCGGCTCGGCTTACCGTCACCGCGACGTCGAGGTGCCGAAGGACCGCGCACCTGGTGC  
ATGACCCGCAAGCCCAGGCTGCCTGACGCCGCCCCACGACCCGAGCGCCGACCGAAAGGAG  
CGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGGCCCCGCCACCCGCCACCG  
CCCGAGGCCACCGACT

*Fig. 12*

**pCMV-pur-attP (SEQ ID NO: 5)**

CTAGAGTCGGGCGGCCGCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTGGAC  
AAACCACAACATAGAACATGCAAGTAAACAAAGTAAACAACAATTGCATTCAATTGCT  
TTATTTGTAACCATATAAGCTGCAATAACAAGTAAACAACAATTGCATTCAATTGCT  
GTTTCAGGTTCAAGGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTCGGCTTCGACTAGTAGTACTGACGGACACACCGAAGCCCCGGC  
GGCAACCCCTCAGCGGATGCCCGGGGCTCACGTTCCCAGGTCAAAGCGGTTTCGGGA  
GTAGTCCCCAAGTGGGTAACCTTGAGTCTCTCAGTGGGGCGTAGGGTCGCCGACAT  
GACACAAGGGGTGTGACCGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCG  
AGCGCAGACTAGTACAAGCGAATTGATCCGTCACCGATGCCCTGAGAGCCTCAACCCAG  
TCAGCTCTTCCGGTGGCGCGGGCATGACTATCGTCGCCGACTTATGACTGCTTCTT  
ATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTTCCGCTTCCGCTCACTGACTCGC  
TGCCTCGGTCTGCTGGCGAGCGGTATCAGCTCACTCAAAGCGGTAAACGGTTA  
TCCACAGAATCAGGGATAACGCAAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG  
GAACCGTAAAAAGGCCGTTGCTGGCTTTCCATAGGCTCCGCCCGTACGAGCAGC  
ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCGACAGGACTATAAGATAACCAGGG  
TTTCCCCCTGGAAGCTCCCTCGTGCCTCTCGTGTGCGCTTACCGGATAACCT  
GTCCGCCTTCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTGTAGGTATCTCA  
GTTCGGTGTAGGTCGTTGCTCCAAGCTGGCTGTGCGACGAAACCCCGTTAGGCCGAC  
CGCTGCCCTTATCCGTAACTATCGTCTTGAGTCCAACCCGTAAGACACGACTTATGCC  
ACTGGCAGCAGCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGGGTGCTACAGAGT  
TCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTG  
CTGAAGCCAGTTACCTCGGAAAAGAGTTGGTAGCTTGTACCGGAAACAAACCC  
TGGTAGCGGTGGTTTTGCAAGCAGATTACGCGCAGAAAAAGGATCTCAAG  
AAGATCCTTGATCTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGG  
ATTGGTCACTGAGATTACAAAAGGATCTCACCTAGATCCTTTAAATTAAAATGAAG  
TTTAAATCAACTAAAGTATATAGATAACTTGGTCTGACAGTTACCAATGCTTAATCA  
GTGAGGCACCTATCTCAGCGATCTGTCTTTCGTTCATCCATAGTGGCTGACTCCCCGTC  
GTGTAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCG  
AGACCCACGCTCACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGAAGGGCGAGC  
GCAGAAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGGTGCCTGGAAAGCT  
AGAGTAAGTAGTCGCCAGTTAATAGTTGCGCAACGTTGGCCATTGCTACAGGCATCGT  
GGTGTACGCTCGTGTGGTATGGCTTCACTCAGCTCCGGTCCCAACGATCAAGGCAG  
TTACATGATCCCCATGGTGTGCAAAAAGCGGTTAGCTCTCGGTCTCCGATCGTGT  
AGAAGTAAGTGGCCGAGTGGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTAC  
TGTGATGCCATCCGTAAGATGCTTCTGTGACTGGTGGTACTCAACCAAGTCATTCTGAG  
AATAGTGTATGCCGCGACCGAGTTGCTCTGCCGGCGTCAATACGGATAATACCGCGCCA  
CATAGCAGAACTTAAAGTGTCTCATCTGGAAAACGTTCTCGGGCGAAAACCTCTCAAG  
GATCTTACCGCTGGAGATCCAGTTGATGTAACCCACTCGGCACCCACTGATCTTAC  
CATCTTTACTTACCGCTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAA  
AAGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATG  
AAGCATTTATCAGGGTATTGCTCATGAGCGGATACATATTGAATGTATTAGAAAATA  
AACAAATAGGGTTCCGCGCACATTCCCGAAAAGTGCACCTGACGCCCTGTAGCG  
GCATTAAGCGCGCGGGTGTGGTGGTACCGCGAGCGTGAACGCTACACTTGCCAGGCC  
AGCGCCCGCTCCCTTCGCTTCTCCCTTCTCGCCACGTTGCCGGCTTCCCGTC  
AAGCTAAATCGGGGCTCCCTTAGGGTCCGATTAGTGTCTTACGGCACCTCGACCC  
AAAAAACTTGATTAGGGTGTGGTACGTAGTGGGCCATGCCCTGATAGACGGTTTC  
CCCTTGACGTTGGAGTCACGTTCTTAATAGTGGACTCTTGTCCAAACTGGAAACAC  
TCAACCTATCTCGGTCTATTCTTGTATTATAAGGGATTGCGATTGCCCTATTGG  
TTAAAAAAATGAGCTGATTAAACAAAATTTAACGCGAATTAAACAAAATATTAAACGTTAC  
AATTTCCTTACGCCATTCAAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCTCT  
CGCTATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATTAAAGGTACGGGAGGTACTT  
GGAGCGGCCGCAATAAAATATCTTATTTCATTACATCTGTGTGTTGGTTTGTGAA  
TCGATAGTACTAACATACGCTCTCCATCAAAACAAAACAAACTAGCAAAATA

GGCTGTCCCCAGTGCAGGTGCCAGAACATTCTATCGATAGGTACCGAGCTCTT  
ACGCGTGCTAGCCTCGAGCAGGAATCTATACTACATTGAATCAATATTGGCAATTAGCCATATTA  
GTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCATTGCATACGTTGTATCTATA  
TCATAATATGTACATTTATATTGGCTCATGCTAACATATGACGCCATGTTGACATTGATTAT  
TGACTAGTTATTAATAGTAATCAATTACGGGTCTAGTTCATAGCCATATATGGAGTC  
CGCGTTACATAACTACGGTAAATGGCCCGCTGGCTGACGCCAACGACCCCGCCCATT  
GACGTCAATAATGACGTATGTTCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAAT  
GGGTGGAGTATTACGGTAAACTGCCACTGGCAGTACATCAAGTGTATCATATGCCAAGT  
CCGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGAC  
CTTACGGGACTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA  
TGCCTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTCCAAGT  
CTCCACCCATTGACGTCAATGGAGTTGTTGGCACCAAAATCAACGGGACTTCAA  
ATGTCGTAACAACCTCGCCCCATTGACGCAAATGGCGGTAGCGTGTACGGTGGAGGTCT  
ATATAAGCAGAGCTCGTTAGTGAACCCTCAGATCGCCTGGAGACGCCATCCACGCTGTTT  
GACCTCCATAGAACAGACCCGGACCGATCCAGCCTCCCTCGAACGCTCGACTCTAGGGCTC  
GAGATCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTGGCCACGCCACCGGGTGGCGCA  
CCATCCCTGACCCACGCCCTGACCCCTCACAGGAGACGCCACCGAGTACAA  
GCCACCGGTGCGCCTCGCACCCGGACGACGTCCCCCGGCGTACGCCACCTCGCCGCG  
CGTCGGCGACTACCCGCCACGCCACCCGTGACCCGGACGCCACATCGAGCGGGTC  
ACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGCTCGACATCGCAAGGTGTGGTC  
GGACGACGGCGCGCGGTGGCGGTCTGGACACGCCGGAGAGCGTCAAGCGGGCGGTGT  
TCGCGAGATCGCCCGCATGGCGAGTGTAGCGGTTCCCGCTGGCGCGCAGCAACAG  
ATGGAAGGCCTCTGGCGCCGACCGGCCAACGGAGCCCGCGTGGTCTGGCCACCGTC  
CGTCTCGCCCGACCACCGGGCAAGGGCTGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGG  
CGGCCGAGCGCGCCGGGTGCCCCCTTCTGGAGAACCTCCGGCCCCCAACCTCCCCTC  
TACGAGCGGCTCGGCTTCACCGTACCGCCACGTGAGGTGCCGAAGGACCGCGCACCTG  
GTGCATGACCCGCAAGCCGGTGCCTGACGCCGCCACGACCCGCAAGGCCACCGAAA  
GGAGCGCACGACCCATGGCTCCGACCGAAGGCCACCCGGGGCCCCGCCACCCGCA  
CGCCCCCGAGGCCACCGACT

**Fig. 13**

**pCMV-EGFP-attB (SEQ ID NO: 6)**

CTAGAGTCGGGGCGGCCGCTTCGAGCAGACATGATAAGATACTTGATGAGTTGGAC  
AAACCACAACATAGAATGCAGTGAAAAAAATGCTTATTGTGAAATTGATGCTATTGCT  
TTATTGTAACCATTATAAGCTGCAATAAACAGTTAACACAACATTGCATTCTATTAT  
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAAACCTCTACAAATGTG  
GTAAAATCGATAAGGATCAATTCCGCTTCAGGTACCGTCGACGATGTAGGTACCGGTCTCGA  
AGCCGCGGTGCGGGTGCAGGGCGTGCCTGGGCTCCCCGGCGTACTCCACCTCACCC  
ATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGCGAACCGCGGGCG  
CACCGGGAAGCCCTCGCCCTCGAACCGCTGGCGCGTGGTCACGGTGGACCGGACGGT  
CGACGGCGTCGGCGGGTGCAGGATAACGCGGGCAGCGTCAGCGGGTCTCGACGGTCACGGCG  
GGCATGTCGACAGCGAACCGTACGCGATGCCCTGAGAGGCCCTCAACCCAGTCAG  
CTCCTCCGGTGGCGCGGGCATGACTATCGTCGCGCACTATGACTGTCTTCTTATCA  
TGCAACTCGTAGGAACAGGTGCCGAGCGCTTCCGCTTCGCTACTGACTCGCTGCG  
CTCGGTCGTTGGCTGCGCGAGCGGTATCAGCTCACTCAAAGGCGTAATACGTTATCCA  
CAGAACATCAGGGATAACCGAGGAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC  
CGTAAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAA  
AAATCGACGCTCAAGTCAGAGGTGGCGAACACCGACAGGACTATAAGATAACCAGGCGTTTC  
CCCCCTGGAAGCTCCCTCGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATAACCTGTC  
GCCTTCTCCCTCGGGAAAGCGTGGCGTTCTCAATGCTCACGCTGTAGGTATCTCAGTC  
GGTAGGTGCTCGCTCCAAGCTGGCTGTGACGAACCCCCCGTCAGCCGACCGCT  
GCGCCTATCCGTAACTATCGCTTGAGTCAACCGGTAAGACACGACTATCGCCACTG  
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGAGGCGGTCTACAGAGTTCTT  
GAAGTGGTGGCTTAACACTACGGCTACACTAGAACAGTATTGGTATCTGCGCTCTGCTGA  
AGCCAGTTACCTCGGAAAAAGAGTTGGTAGCTTGTGACAGTTACCAATGCTTAATCAGTGA  
GGCACCTATCTCAGCGATCTGCTATTGTTCATAGTGCCTGACTCCCGTCGTGT  
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATAACCGCGAGAC  
CCACGCTCACCGGCTCCAGATTATCAGCAATAACAGCCAGGCCAGGGCCAGCGCAG  
AAGTGGCCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCATTGCTACAGGATCGTGGT  
TAAGTAGTTCGCCAGTTAATAGTTGCGCAACGTTGCTTGCCTACAGGATCGTGGT  
TCACGCTCGTGTGGTATGGCTCATTCACTCCGGTTCCCAACGATCAAGGCGAGTTAC  
ATGATCCCCCATGTTGCAAAAAGCGGTAGCTCTCGGCTCCGATCGTGTGTCAGAA  
GTAAGTGGCCGAGTGTATCTACTCATGGTTATGGCAGCACTGCATAATTCTTACTGTC  
ATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGACTCAACCAAGTCATTCTGAGAATA  
GTGTATGCCGCGACCGGAGTTGCTTGCCTTGCCTTGCCTACAGGATAATACCGGCCACATA  
GCAGAACTTTAAAGTGTCTCATATTGAAAACGTTCTCGGGCGAAAACCTCAAGGATC  
TTACCGCTGTTGAGATCCAGTTGCTGTAACCCACTCGTCGACCCAACTGATCTCAGC  
TTTACTTACCCAGCGTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCCAAAAAGG  
GAATAAGGGCGACACGAAATGTTGAATACTCATACTCTTCCCTTTCAATATTATGAAAGC  
ATTTATCAGGGTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAATAAAC  
AAATAGGGTTCCCGCAGCATTTCCCGAAAAGTGCACCTGACGCGCCCTGTAGCGCGCAT  
TAAGCGGGCGGGTGTGGTTACGCGCAGCGTACCGCTACACTTGCCAGCGCCCTAGCG  
CCCGCTCTTCGCTTCTCCCTTCTCGCCACGTTCGCCGGCTTCCCGTCAAGC  
TCTAAATCGGGGCTCCCTTAGGGTCCGATTTAGTGTCTACGGCACCTCGACCCAAA  
AACTTGATTAGGGTGTGGTTACCGTAGTGGGCCATGCCCTGATAGACGGTTTCGCCC  
TTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCTGTTCCAAACTGGAACAAACTCAA  
CCCTATCTCGGTCTATTCTTTGATTATAAGGATTGCGATTCCGGCTATTGGTTAA  
AAAATGAGCTGATTTAACAAAATTTAACGCGAATTAAACAAATATTACGTTACAATT  
TCCCATTGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTTCGCT  
ATTACGCCAGCCAAAGCTACCATGATAAGTAAGTAATATTAGGTACGGGAGGTACTTGGAG  
CGGCCGAATAAAATCTTATTTCATTACATCTGTGTGGTTTTGTGTAATCGA

TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAACAAACTAGCAAAATAGGCT  
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTCTATCGATAGGTACCGAGCTTACGC  
GTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGCAATTAGCCATATTAGTC  
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCATTGCATACGTGTATCTATATCAT  
AATATGTACATTATATTGGCTATGTCAAATATGACCGCCATGTTGACATTGATTATTGAC  
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTCATAGCCCATAATGGAGTTCCGCG  
TTACATAACTACGGTAAATGGCCCGCTGGTGAACGCCAACGACCCCCGCCATTGACG  
TCAATAATGACGTATGTTCCCATAAGTAACGCCAATAGGGACTTCCATTGACGTCATGGGT  
GGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC  
CCCCTATTGACGTCATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTA  
CGGGACTTCCACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCG  
GTTTGCGAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGGATTCCAAGTCTCC  
ACCCCATTGACGTCATGGGAGTTGTTTGGCACCAAAATCAACGGACTTCCAAAATGT  
CGTAACAACACTCCGCCCCATTGACGCAAATGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT  
AAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTTTGACC  
TCCATAGAACAGACACGGGACCGATCCAGCCTCCCTCGAACGCTCGACTCTAGGGCTCGAGA  
TCCCCGGGTACGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGC  
CCATCCTGGTCGAGCTGGACGGCAGCTAAACGCCACAAGTCAGCGTGTCCGGCGAGGGC  
GAGGGCGATGCCACCTACCGCAAGCTGACCTGAAGTTCATCTGCACCACCGGCAAGCTGCC  
CGTGCCTGGCCACCCCTCGTACCGCACCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACC  
CCGACCACATGAAGCAGCACGACTTCAAGTCGCCATGCCGAAGGCTACGTCCAGGAG  
CGCACCATCTTCTCAAGGACGACGGCAACTACAAGACCCGCCGAGGTGAAGTTGAGGG  
CGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC  
TGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGCTATATCATGGCGACAAGCAG  
AAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT  
CGCCGACCACTACCAGCAGAACACCCCCATGGCGACGGCCCGTGTGCTGCCGACAACC  
ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC  
CTGCTGGAGTTCGTGACGCCGCCGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAG  
CGGCCGCTCGAGCATGCAT

***Fig. 14***

**p-12.0-lys-LSP1FNMM-CMV-pur-attB (SEQ ID NO: 7)**

GGGCTGCAGGAATTGCGATTGCCGCTTCTTGATATTCACTCTGTTGATTTCATCTCTTCT  
TGCGATGAAAGGATAAACAGTCGTATAACAGTCGTGAGGAAACTTGGTATTTCTTC  
TGATCAGTGTGTTTATAAGTAATGTTGAATATTGGATAAGGCTGTCCTTGTCTGGG  
AGACAAAGCCCACAGCAGGTGGTGGGTGGCAGCTCAGTGACAGGAGAGGTTTT  
TTGCCTGTTTTTTTTTTTTTTAAGTAAGGTGTTCTTTCTTAGTAAATT  
CTACTGGACTGTATGTTGACAGGTAGAAAACATTCTCAAAGAAGAACCTTTGGAAA  
CTGTACAGCCCTTTCTTCATTCCCTTTGCTTCTGTGCCAATGCCTTGGTTCTGATT  
GCATTATGGAAAACGTTGATCGGAACCTGAGGTTTATTTATAGTGTGGCTTGAAAGCTG  
GATAGCTGTTGTTACACGAGATACCTTATTAAGTTAGGCCAGCTGATGCTTATTTTC  
CCTTGAAAGTAGTGAGCGTTCTGGTTTTCCCTTGAAACTGGTGAGGCTTAGATTTT  
CTAATGGGATTTTACCTGATGATCTAGTTGCATACCCAAATGCTTGAAATGTTTCCA  
GTTAACATGTTGATAACTCGGATTACATGTTGATATACTTGTGATCTGTGTTCTAGTA  
AAAATATATGGCATTATAGAAATACGTAATTCTGATTCCCTTCTTATCTCTATGCT  
CTGTTGTCAGGTCAAACAGACTTCACTCTATTATTTATAGAATTATATGCAAGTC  
TGCGTTGGTCTTGTTGTAAGGATACGCCCTAAATTCTAGAGCGATGCTCAGTAAG  
GCGGGTTGTCACATGGGTCAAATGTAACACGGGCAGCTGGCTGCTGCCTCCGAGATC  
CAGGACACTAAACTGCTCTGCACTGAGGTATAAATCGCTCAGATCCCAGGGAAGTGCAGA  
TCCACGGCATATTCTTAAAGAAGAATGAAACTTTCTAAATTATTTGGCATAGGAAGCAA  
GCTGCATGGATTGTTGGGACTTAAATTATTTGGTAACGGAGTGCATAGGTTAAACAC  
AGTTGCAGCATGCTAACGAGTCACAGCGTTATGCAGAAGTGTGCGCTGGATGCCCTGCG  
GCTGTTACGGCACTGCCTGCACTGAGCATTGCAAGATAGGGTGGGGTGCCTTGTGCG  
TTCCACACGCTGCCACACGCCACCTCCCGAACACATCTCACCTGCTGGTACTTTCAA  
ACCATCTTAGCAGTAGTAGATGAGTTACTATGAAACAGAGAAGTCTCAGTTGGATATTCT  
CATGGGATGTCCTTTCCATGTTGGCAAAGTATGATAAAAGCATCTATTGTAATT  
TGCACTTGTTAGTCTCTGAATCCTTCTATAGCACCCTATTGAGCAGGTGAGGCTCTG  
GTGTGGCCTGTCCTGCTCAATCTTAAAGCTCTTGGAAATACTGACTGACTTGATIG  
AAGTCTCTGAAAGATAGTAAACAGTACTTACCTTGATCCCAATGAAATCGAGCATTCA  
TGTAAAAGAATTCCGCCTATTCAACCATGTAATGTAATTTCACCCCCAGTGCACACT  
TTGGAATATATTCAAGTAATAGACTTGGCCTCACCCCTTGTGACTGTATTGTAATAG  
AAAATATTAAACTGTGCATATGATTATTACATTATGAAAGAGACATTGCTGATCTCA  
AATGTAAGAAAATGAGGAGTGCCTGCTTATAAATACAAGTGATTGCAAATTAGTGCAG  
GTGCTCTAAAAAAGTAATATAAAAAGGACCAGGTGTTACAAGTGAAT  
ACATT CCTATTGGTAAACAGTTACATTGATCCCAATGAAATCGAGCATT  
ACATAAGGCTGTATTGCTTCTGTACCTGCATTGCTTCCATTCCCAATTGCAACAAGGAT  
GTCTGGTAAACTATTCAAGAAATGGCTTGAAATACAGCATGGAGCTGCTGAGTTGGA  
ATGCAGAGTTGCACTGCAAATGTCAAGGAAATGGATGTCTCTCAGAATGCCA  
GATTTATATGTGTATAGTAAGCAGTTCTGATTCCAGCAGGCCAAAGAGTCTGCTGAA  
TGTTGTGTTGCCGGAGACCTGTATTCTCAACAAGGTAAGATGGTATCCTAGCAACTGCG  
TTTAAATACATTTCAGCAGAAGTACTTAGTTAATCTACCTTGGATGTTCATCAT  
TTTAGATGTTACTTGAAATACTGCATAACTTTAGCTTCTATGGGTTCTTTTTCTAG  
CCTTTAGGAGACTGTTAAGCAATTGCTGTCACCTTTGTTGGCTTAAACTGCAATAG  
TAGTTACCTGTATTGAAAGAAATAAGACCATTTTATATTAAAAAAACTTTGTCTGTC  
TTCATTGACTGTCGATATCCTTGCAGTGCCCATTATGTCAGTTCTGTCAGATATT  
ACATCAAAACTTAACGTGAGCTCAGTGGAGTTACAGCTGCCTTGTGCTGTTATT  
CTGAAACTAGAAATGATGTTGTCCTCATCTGTCATCAAACACTCATGCAGAGTGT  
TAGTGAGAAATGCAACATTGAAACTTTAAAGTCACATTGATATTGAGGTT  
TTCATTGGAAATATATTGTTCTAGACTGCATAGCTCTGAAATGCAAGTCTGAT  
TGGCATGAAGAACAGCACAGCCTTCAACTTAACAGACAGTGCCTCAGGAGAAAGT  
AAGCAAGGGCACAGGTCCATGAAATAGAGACAGTGCCTCAGGAGAAAGTGAAC  
CTTGGCTAGTGTCTAAATCTGAGTGTAGGAAAGTAACACCCGATTCTTGAAAGGG  
AGCTTAAATGCTCCAAATTGAAGGTGGCAGGCAACTGGCACTGGTATT  
TGTCTCAGTTCGCAGCTAACCTGGCTTCTCCACTATTGAGCAGTGGACTAT  
AGCCTGGCTGGACTGTGGCTGGCTGGGGACTGTG  
AGAGGCCAGGTGAAGGTTGGATGGGTGAAAGGAGTGCCTGGCTGGGGACTGTG

GGGACTCCAAGCTGAGCTGGGTGGGCAGCACAGGGAAAAGTGTGGGTAACATTTTAAG  
TACTGTGTTGCAAACGTCATCTGCAAATACGTAGGGTGTACTCTGAAGATTAACAGT  
GTGGGTCAGTAATATATGGATGAATTACAGTGGAAAGCATTCAAGGGTAGATCATCTAACG  
ACACCAGATCATCAAGCTATGATTGGAAAGCGGTATCAGAAGAGCGAGGAAGGTAAGCAGTCT  
TCATATGTTCCCTCACGTAAGCAGTCTGGAAAGTAGCAGGCCAGAGCCTGTGGTAGCTGGA  
GAAATAATTCAAGGAGCATGTGCTAGGAGAACTTCTGCTGAATTCTACTTGCAAGAGCTTT  
GATGCCTGGCTTGGTGCCTCTGCAGCACCTGCAAGGCCAGAGCCTGTGGTAGCTGGA  
GGGAAAGATTCTGCTCAAGCTCAAGCTTCAGCAGGTATTGCTTGTCTTCCCCAGCA  
CTGTGCAGCAGAGTGGAACTGATGTCAGACGCCCTGTCCACTACCTGTTGCTGCAGGCAGA  
CTGCTCTAGAAAAAGAGAGCTAACACTATGCCATAGTCTGAAGGTAACATGGGTTTAAAA  
AAGAAAACACAAAGGCAAAACGGCTGCCCATGAGAAGAAAGCAGTGGTAAACATGGTAGA  
AAAGGTGCAGAAGGCCAGGCAGTGTGACAGGCCCTCTGCCACCTAGAGGCCAGAACAA  
GCTTCCTGCCTAGGGCTCTGCCCGGAAGTGCCTGGTTCTTGGTGGGTTTGGTGGCGT  
TTGGTTTGAGATTAGACACAAGGAAGCCTGAAAGGGAGGTGTTGGGCACTATTGGTT  
GTAAAGCCTGTACTTCAAATATATATTGTGAGGGAGTGTAGCGAATTGGCAATTAAAA  
TAAAGTGCAAGAGATTGAGGCTGAGTAGTTGAGAGGGTAACACGTTAATGAGATCTTCT  
GAAACTACTGCTCTAAACACTTGGTTGAGTGGTGAGACCTGGTAGGGTCTTGGCT  
TACATGTCATGCACTTGTCTTCCATCCACATCCATGCATCCACATCCACGCA  
TTTGTCACTTATCCCATATCTGTCATATCTGACATACCTGTCTTGTCACTTGGTCAGAA  
GAAACAGATGTGATAATCCCAGCCGCCAAGTTGAGAAGATGGCAGTTGCTCTTCCC  
TTTGTCTGCTAAGTAAGGATTCTCCTGGCTTGACACCTCACGAAATAGTCTTCC  
TTACATTCTGGCATTATTCAAATATCTTGGAGTGCCTGCTCTCAAGTTGTCTTCC  
TACTCTTAGAGTGAATGCTCTTAGAGTGAAGAGAAGAGAAGAGATGTTGGCCAGTTC  
TCTGATGAACACACCTCTGAATAATGGCCAAGGTGGGTTCTGAGGAACGGCAG  
CGTTTGCCTCTGAAAGCAAGGAGCTGCGGAGTTGCAGTTATTGCAACTGATGGTGGAA  
CTGGTGCCTAAAGCAGATTCCCTAGGTTCCCTGCTACTTCTTCTTCTTGGCAGTCAGTT  
TATTCTGACAGACAAACAGCCACCCCCACTGCAGGCTTAGAAAGTATGTGGCTCTGCC  
GTGTGTTACAGCTCTGCCCTGGTGAAAGGGATTAAAAGGGCACCATTCACTCCAAACAGG  
ATCCTCATTGATCAAGCTGTAAGGAACCTGGCTCCAACCTCAAAACATTAAATTGGAG  
TACGAATGTAATTAAAACGTCATTCTGCATTCTAAGTCATTAGTCTGGACTCTGCAGCA  
TGTAGGTGGCAGCTCCACTTCTCAAAGACCACTGATGGAGGAGTAGTAAAAATGGAGAC  
CGATTCAAAACCAACGGAGTGTGCGGAAGAAACTGATGGAAATAGTCATGAATTGTG  
TGGTGGACATTAAATACATAAAACTACTTCAATGAGGTGGAGAAGGTCAAGTGGT  
TATTAGCAGCCATAAAACCAAGGTGAGCGAGTACCATTTCTCTACAAAGAAAAACGATTCTG  
AGCTCTGCGTAAGTATAAGTTCTCCATAGCGGCTGAAGCTCCCCCTGGCTGCCATCT  
CAGCTGGAGTGCAGTGCCTTCTGGGTTCTCCTACAGCAGTAATGGACAATACTTC  
ACAAAAATTCTTCTTCTCTGTCATGTGGATCCCTACTGTGCCCTCTGGTTTACGTTA  
CCCCCTGACTGTTCCATTCAAGGGTTTGGAAAAGAGAAAAAGAATTGGAAATAAAACATGTC  
TACGTTATCACCTCTCCAGCATTGGTTTAATTATGTCATAACTGGCTTAGATTGG  
AAATGAGAGGGGGTTGGGTGATTACCGAGGAACAAAGGAAGGCTTATATAAAACTCAAGTCT  
TTTATTAGAGAACTGGCAAGCTGTCAAAACAAAAGGCCTTACCAACAAATTAGTGAAT  
AGCCGCTATAGCCAGCAGGGCCAGCACGAGGGATGGTGCAGTGGCACTATGCCACGCC  
TGCTTGTGACTCTGAGAGCAACTGCTTGGAAATGACAGCAGTCAGTGGCAATTCTTGT  
TCAGAATGCGTAGAGCGTGTGCTGGCGACAGTTCTAGTTAGGCCACTTCTTTCT  
TCTCTCTCATTCTCTAACGATGTCTCCATGCTGGTAATCCAGTCAGTGAACGTTCAA  
CAATGAATCCATCACTGTAGGATTCTGTTGATCAAATCTTGTGAGGTCTATAAAAT  
ATGGAAGCTTATTATTTCGTTCTTCATATCAGTCTCTATGACAATTACATCCAC  
CACAGCAAATTAAAGGTGAAGGAGGCTGGTGGATGAAGAGGGCTTCTAGCTTACGTT  
TCCCTGCAAGGCCACAGGAAAATGCTGAGAGCTGAGAATACAGCCTGGGTAAGAAGTCA  
GTCTCTGCTGGACAGCTAACCGCATCTTATAACCCCTCTGAGACTCATCTTAGGACCAA  
ATAGGGTCTATCTGGGTTTTGTTCTGCTGTTCTCTGGAAAGGCTATCTCACTATTCA  
CTGCTCCACGGTTACAAACCAAGATAACAGCCTGAATTCTAGGCCACATTACATAAA  
TTTGACCTGGTACCAATTGTTCTCTATATAGTTATTCTCCCACTGTGTTAACCC  
TTAAGGCATTCAAGAACACTAGAATGTTGGATTGGAAGGGCCTAAACATC

ATCCATTTCCAACCCTCTGCCATGGGCTGCCACCCACTGGCTCAGGCTGCCAGGGCC  
CCATCCAGCCTGGCCTTGAGCACCTCCAGGGATGGGCACCCACAGCTCTCTGGCAGCCT  
GTGCCAACACCTCACCACCTCTGGTAAAGAATTCTCTTTAACATCTAAATCTAAATCTCT  
TCTCTTTAGTTAAAGCATTCCCTTTCCCGTGTCTATCTGTCCAAGAAATGTGTATT  
GGTCTCCCTCTGTTATAAGCAGGAAGTACTGGAAGGCTGCAGTGAGGTCTCCCCACAGCC  
TTCTCTCTCCAGGCTGAACAAGCCCAGCTCCTCAGCCTGTCTCGTAGGAGATCATCTA  
GTGGCCCTCCTCTGGACCCATTCCAACAGTCCACGGCTTCTGTGGAGCCAGGTCTGG  
ATGCAGTACTTCAGATGGGCCTTACAAAGGCAGAGCAGATGGGACAATCGCTACCCCTC  
CCTGCTGGCTGCCCTGTTGATGCAGCCAGGGTACTGTTGGCCTTCAGGCTCCCAGAC  
CCCTTGCTGATTGTGTCAAGCTTTCATCCACCAGAACCCACGCTCCTGGTTAAACTTC  
TGCCCTCACTTCTGTAAGCTTGTTCAGGAGACTTCATTCTTCTAGGACAGACTGTGTTACA  
CCTACCTGCCCTATTCTGCATATATACATTCACTGTTCTGTAAACAGGACAGAAT  
ATGTATTCTCTAAACAAAATACATGCAGAATTCTAGTGCATCTCAGTAGGGTTTCATG  
GCAGTATTAGCACATAGTCAATTGCTGCAAGTACCTTCAAGCTGCGGCCTCCATAATC  
CTGTATTGGGATCAGTTACCTTGGGTAAGCTTGTATCTGCAGAGACCCCTGGGGTT  
CTGATGTGCTTCAGCTGCTCTGTTCTGACTGCACCATTCTAGATCACCCAGTTGTTCC  
TGTACAACCTCCTTGTCCATCCTTCCAGCTTGTATCTTGACAAATACAGGCCTATT  
TTTGTGTTGCTCAGCAGCATTAAATTCTCAGTGTATCTTGTTCTGTTGATGCCACTG  
GAACAGGATTTCAGCAGTCTGCAAAGAACATCTAGCTGAAACTTCTGCCATTCAATAT  
TCTTACAGTCTTCTTGTGAGGTGAGCCATAAAATTACTAGAACCTCGTCACTGACAAGT  
TTATGCATTTATTACTCTATTATGTACTACTTGAACATAACACAGACACGACATATT  
TGCTGGGATTTCACAGTGTCTCTGTCCTCACATGGTTTACTGTCATACTCCGTTAT  
AACCTTGGCAATCTGCCAGCTGCCATCACAAGAAAAGAGATTCTTTTATTACTTCTC  
TTCAGCCAATAACAAAATGTGAGAAGGCCAAACAAGAACCTGTGGGCAGGGCTGCCATCAA  
GGGAGAGACAGCTGAAGGGTTGTAGCTCAATAGAATTAAAGAAATAAAAGCTGTGTCAG  
ACAGTTTGCCTGATTATACAGGCACGCCAACAGCCAGAGAGGCTGCTGCCAAGGCCACC  
TTGCAGTCCTGTTGTTGAAAGATAAGTCAGGTAACCTTCTGGTAATTGCGTGGAGAAT  
CATGATGGCAGTCTGCTGTTACTATGGTAAGATGCTAAACATTAGGAGACAGCAAAGTAAC  
ACTGCTGCTGTAGGTGCTGCTGCTATCCAGACAGCGATGGCAGTCACACCAAGATGAGGG  
ATGCTCCAGCTGACGGATGCTGGGCAGTAACAGTGGTCCATGCTGCCTGCTCATTAGC  
ATCACCTCAGCCCTCACCAGCCATCAGAAGGATCATCCCAAGCTGAGGAAAGTGCTCATC  
TTCTTCACATCATCAAACCTTGGCCTGACTGATGCCCTCCGGATGCTAAATGTGGTCACT  
GACATCTTATTCTATGATTCAAGTCAGAACCTCCGGATCAGGAGGGAACACATAGTG  
GGAATGTACCCCTCAGCTCCAAGGCCAGATCTCCTCAATGATCATGCTACTTAGGAA  
GGTGTGTTGTAATGAGATTGCTTCTGGCATGGAGGAGTTGTACACTTGCAA  
TTGAATACCAGAGAAAAAGAAAAGCTCTGTTATTTTCTTCTGCTGTAGGAACATT  
AATAAAGGATGCAGTCCAAATGTCATAATCTCAGGGTCTGAAGGAGGATCAGAAACTGTG  
TATACAATTTCAGGCTTCTGAAATGCAGCTTGTAAAGCTGTTCTGGCGAGGAGTACT  
AGTCAGAACCTCGGAAACAGGAACAAATGTCCTCAAGGTGCAGCAGGAGGAAACACCTTG  
CCATCATGAAAGTGAATAACCAACTGCGCTGAAGGAATCCAGCTCCTGTTGAGCAGGTGCT  
GCACACTCCCACACTGAAACACAGTTCATTTATAGGACTCCAGGAAGGATCTTCTCT  
TAAGCTCTTAATTATGGTACATCTCCAGTGGCAGATGACTATGACTACTGACAGGAGAAT  
GAGGAACTAGCTGGAAATATTCTGTTGACCACCATGGAGTCACCCATTCTTACTGGTA  
TTTGGAAATAATAATTCTGAAATTGCAAAGCAGGAGTTAGCGAAGATCTCATTCTTCCATG  
TTGGTGACAGCACAGTTCTGGCTATGAAAGTCTGCTTACAAGGAAGAGGATAAAATCATAG  
GGATAATAAAATCTAAGTTGAAGACAATGAGGTTTAGCTGCATTGACATGAAGAAATTGA  
GACCTCTACTGGATAGCTATGGTATTACGTTGCTTTGCTAGTTACTTATTGACCCAG  
CTGAGGTCAAGTATGAACTCAGGCTCTCGGGCTACTGGCATGGATTGATTACATACAAC  
TAATTTAGCAGTGTAGTTAGGGTTATGAGTACTTTGCAAGTAAATCATAGGTTAGTAATG  
TTAATCTCAGGGAAAAAAAGCCAACCCCTGACAGACATCCCAGCTCAGGTGGAAATC  
AAGGATCACAGCTCAGTGCAGGCTCCAGAGAACACAGGGACTCTCTTAGGACCTTATGT  
ACAGGGCCTCAAGATAACTGATGTAGTCAGAAGACTTCCATTCTGGCCACAGTTCAGCTG  
AGGCAATCCTGGAATTCTCTCCGCTGCACAGTTCCAGTCATCCCAGTTGTACAGTTCTG  
GCACTTTGGGTCAAGGCCGTGATCCAAGGAGCAGAACAGTCCAGCTATGGTCAGGGAGTGCC

TGACCGTCCCAACTCACTGCACTCAAACAAAGGCGAAACCACAAGAGTGGCTTTGTTGAAA  
TTGCAGTGTGGCCCAGAGGGCTGCACCAGTACTGGATTGACCACGAGGCAACATTAATCCT  
CAGCAAGTGCAATTGCAGCCATTAAATTGAACTAACGTGATACTACAATGCAATCAGTATCA  
ACAAGTGGTTGGCTTGGAAAGATGGAGTCTAGGGGCTCTACAGGAGTAGCTACTCTTAATG  
GAGTTGCATTTGAAGCAGGACACTGTGAAAAGCTGGCCTCTAAAGAGGCTGCTAAACATT  
AGGGTCAATTTCAGTGCACTTCTGAAGTGTCTGCAGTCTCCCCTGCAAAGCTGCCAAA  
CATAGCACTCCAATTGAATACAATTATATGCAGGCGTACTGCTTCTGCCAGCACTGCTCT  
TCTCAAATGAACACTCAAACAAATTCAAAGTCTAGTAGAAAGTAACAAGCTTGAATGTCA  
TTAAAAAGTATATCTGCTTCAGTAGTTAGCTTATTTATGCCACTAGAAACATCTGTAC  
AAGCTGAACACTGGGCTCCAGATTAGTGGTAAAACCTACTTTATACATAGAATCATA  
GAATGGCCTGGGTTGGAAGGGACCCCAAGGATCATGAAGATCCAACACCCCCGCCACAGGCA  
GGGCCACCAACCTCCAGATCTGGTACTAGACCAGGCCAGGGCTCCATCCAACCTGGCC  
ATGAACACACTCCAGGGATGGGACATCCACAAACCTCTGGGAGCCTGTGCCAGCACCTCAC  
CACCTCTCTGTGAAGAACCTTTCCCTGACATCCAATCTAACGCTTCCCTCCTGAGGTTAG  
ATCCACTCCCCCTGTGCTATCACTGTCTACTCTTGAAAAGTTGATTCTCCTCCTTTG  
GAAGGTTGCAATGAGGTCTCCTTGAGCCTCTTCTCTGAGGATGAACAAGCCCAGCT  
CCCTCAGCCTGTCTTATAGGAGAGGTTGCTCCAGCCCTCTGATCATCTTGAGGCCCTCCTC  
TGGACCCGCTCCAAGAGCTCCACATCTTCTGTACTGGGGGCCCCAGGCCTGAATGCAGTA  
CTCCAGATGGGGCCTCAAAAGAGCAGAGTAAAGAGGGACAATCACCTCCTCACCCCTGCTGG  
CCAGCCTCTCTGATGGAGCCCTGGATACAACACTGGCTTCTGAGGCTGCAACTTCTCCTTAT  
CAGTTCCACTATTAAAACAGGAACAATACAACAGGTGCTGATGGCCAGTGCAGAGTTTCA  
CACTTCTTCATT CGGTAGATCTTAGATGAGGAACGTTGAAGTTGTGCTTCTGCGTGTGCTT  
CTTCCCTCCTCAAATACTCCTGCCTGATACCTCACCCACCTGCCACTGAATGGCTCATGGC  
CCCCTGCA GCCAGGGCCCTGATGAACCCGGCACTGCTTCAGATGCTGTTAATAGCACAGTA  
TGACCAAGTTGCACCTATGAATAACACAAACAATGTGTTGCACTCCTCAGCACTTGAGAAGAA  
GAGCCAAATTGCAATTGTCAGGAATGGTTAGTAATTCTGCCAATTAAAACCTTGTTATCT  
ACCATGGCTGTTTATGGCTGTTAGTAGTGGTACACTGATGATGAACAATGGCTATGCAGT  
AAAATCAAGACTGTAGATATTGCAACAGACTATAAAATTCCCTGTGGCTTAGCCAATGTGG  
TACTTCCCACATGTATAAGAAATTGGCAAGTTAGAGCAATGTTGAAGTGTGGAAAT  
TTCTGTATACTCAAGAGGGCTTTTGACAACACTGTAGAACAGAGGAATCAAAGGGGGTGGG  
AGGAAGTTAAAAGAAGAGGGCAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTACGACACTG  
GCAACATGAGGTCTTGCTAATCTGGTCTTGCTTCCTGCCCTGGCTGCCTAGGGTGC  
GATCTGCTCAGACCCACAGCCTGGCAGCAGGAGGACCTGATGCTGCTGGCTCAGATGAG  
GAGAACATCAGCCTGTTAGCTGCCCTGAAGGATAGGCACGATTGGCTTCCCTCAAGAGGAGT  
TTGGCAACAGTTCAAGAAGGCTGAGACCATCCCTGCTGCTGCAAGGAGATGATCCAGCAGATC  
TTAACCTGTTAGCACCAGGATAGCAGCGCTGCTGGGATGAGACCCCTGCTGGATAAGTT  
TTACACCGAGCTGCTTACCGCAGCTGAACGATCTGGAGGCTTGCCTGATCCAGGGCTGGCG  
TGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGGCTGTGAGGAAGTACTTCAGAGG  
ATCACCCCTGTAACCTGAAGGAGAAGTACAGCCCCCTGCCTGGAAAGTCGTGAGGGCTGA  
GATCATGAGGAGCTTAGCCTGAGCACCAACCTGCAAGAGAGCTTGAGGTCTAAGGAGTAA  
AAGTCTAGAGTCGGGGCGGCCGGCGCTCGAGCAGACATGATAAGATACATTGATGAGTT  
GGACAAACCACAACATAGAATGCAGTAAAAAAATGCTTATTGTGAATTGTGATGCTAT  
TGCTTATTGTAACCATTATAAGCTGCAATAACAAAGTTAACACAAACATTGCAATT  
TTATGTTTCAAGGTTCAAGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAACACAAATTGCAATT  
TGTGGTAAAATCGATAAGGATCCGTCGACCGATGCCCTGAGAGCCTCAACCCAGTCAGCT  
CCTTCCGGTGGCGCGGGCATGACTATCGTCGCCACTTATGACTGTCTTCTTATCATG  
CAACTCGTAGGACAGGTGCCGGCAGCGCTCTCCGCTTCTGCTCACTGACTCGCTGCCT  
CGGTGTTCGGCTGCCGGAGCGGTATCAGCTCACTCAAAGGCGGTAAACCGGTTATCCACA  
GAATCAGGGATAACGCAGGAAAGAACATGTGAGCAGAAAAGGCCAGCAAAAGGCCAGGAACCG  
AAAAAGGCCGCGTTGCTGGCTTTCCATAGGCTCCGCCCCCTGACGAGCATTACAAA  
ATCGACGCTCAAGTCAGAGGTGCCGAAACCCGACAGGACTATAAGATACCAAGGGCTTCCC  
CCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGC  
CTTCTCCCTTCCGGAAAGCGTGGCGTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCG  
TGTAGGTCGTTCGCTCCAAGCTGGCTGTGCAACCCCCCGTTAGCCGACCGCTGC

GCCTTATCCGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGC  
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCCTGCTACAGAGTTCTGA  
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCCTGCTGAAG  
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTTGTACCGGCAAACAAACCACCGCTGGTAG  
CGGTGGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATC  
CTTGATCTTCTACGGGTCTGACGCTCAGGAAAGAAAACACTCACGTTAACGGATTGG  
GTCATGAGATTATCAAAAAGGATCTCACCTAGATCCTTAAATTAAAATGAAGTTAA  
ATCAATCTAAAGTATATGAGTAAACTTGGTCTGACAGTACCAATGCTTAATCAGTGAGG  
CACCTATCTCAGCGATCTGTCTATTGTTCATCCATAGTGCCTGACTCCCCGTGTTAG  
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCCTGCAATGATACCGCGAGACCC  
ACGCTACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGGAGGGCCAGCGCAGAA  
GTGGTCTGCAACTTATCCGCCTCCATCCAGTCTATTAAATTGTTGCCGGAGCTAGAGTA  
AGTAGTTGCCAGTTAATAGTTGCGCAACGTTGCTGACTACAGGCACTGTGGTAG  
ACGCTCGTGTGGTATGGCTTACAGCTCCGGTCCAACGATCAAGGCAGTTACAT  
GATCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCGGTCTCCGATCGTTGTCAGAAGT  
AAGTTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTACTGT  
GCCATCCGTAAGATGCTTTCTGTGACTGGTGGTACTCAACCAAGTCATTCTGAGAATAGT  
GTATCGGGCGACCGAGTTGCTCTGCCGGTCAATACGGATAATACCGGCCACATAGC  
AGAACTTAAAAGTGCATCATTGGAAAACGTTCTCGGGCGAAAACCTCAAGGATCTT  
ACCGCTGTTGAGATCCAGTTGATGTAACCCACTCGTGCACCCAACACTGATCTTCAGCATCTT  
TTACTTTACCAGCGTTCTGGTGAGCAAAACAGGAAGGCAAAATGCCGAAAAAGGGA  
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTCCTTTCAATATTATGAGCAT  
TTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAATAACAAA  
TAGGGGTTCCGCCACATTCCCCGAAAAGTGCACCTGACGCCCTGTAGCGCGCATTA  
AGCGCGCCGGGTGTGGTGTACGCGCAGCGTGAACCGCTACACTTGCCAGCGCCCTAGCGCC  
CGCTCCCTTCGCTTCTTCCCTCCTTCGCCCACGTTGCGCCGGTTCCCGTCAAGCTC  
TAAATCGGGGCTCCCTTACGGGTCCGATTTAGTGTCTTACGGCACCTCGACCCAAAAAA  
CTTGATTAGGGTGTGGTTCACGTAGTGGGCATCGCCCTGATAGACGGTTTCGCCCTT  
GACGTTGGAGTCCACGTTCTTAATAGTGGACTCTTGTCCAAACTGGAACAAACTCAACC  
CTATCTCGGTCTATTCTTGTATTATAAGGGATTGCGATTGCGCTATTGGTAA  
AATGAGCTGATTAAACAAAATTAAACGGAATTAAACAAAATAACGTTACAATTTC  
CCATTGCCATTAGGCTCGCAACTGTTGGGAAGGGCGATCGGTGCGGCCCTTCGCTAT  
TACGCCAGCCAAAGCTACCATGATAAGTAAGTAATATTAAAGGTACGGGAGGTACTGGAGCG  
GCCGCTCTAGAAACTAGTGGATCCCCCGGCCAATAAAATATCTTATTTCATACATCTG  
TGTGTTGGTTTTGTGTAATGATAGTACTAACATACGCTCTCCATCAAAACAAAAGGA  
ACAAAACAAACTAGCAAAATAGGCTGCCCAGTGCAGGTGCAGAACATTCT  
ATCGATAGGTACCGAGCTTACGCGTCTAGCCCTCGAGCAGGATCTATACATTGAATCAA  
TATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATTGGCTATTGGC  
ATTGCATACGTTGTATCTATATCATAATTAGTACATTATATTGGCTCATGTCAAATGAC  
CGCCATGTTGACATTGATTGACTAGTTATTAAATAGTAATCAATTACGGGTCTTACG  
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACC  
GCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTCCCCTAGTAACGCCAATAG  
GGACTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCACTGGCAGTACAT  
CAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTG  
GCATTATGCCAGTACATGACCTTACGGGACTTCTACTTGGCAGTACATCTACGTATTAG  
TCATCGCTATTACCATGGTGATGCGGTTGGCAGTACATCAATGGCGTGGATAGCGGTT  
GACTCACGGGATTCCAAGTCTCCACCCATTGACGTCAATGGAGTTGTTGGCACCA  
AAATCAACGGGACTTCCAAAATGCGTAACAACCTCCGCCCCATTGACGCAAATGGCGGT  
GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGG  
AGACGCCATCCACGCTGTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCTC  
GAAGCTCGACTCTAGGGCTCGAGATCTGCGATCTAAGTAAGCTTGCAATGCCACTGGCAGT  
CCGCCACGACCGGTGCCACCATTCCCTGACCCACGCCCTGACCCCTCACAAGGAGACG  
ACCTTCCATGACCGAGTACAAGCCACGGTGCCTGCCACCCCGCAGCAGTCCCCGGG  
CCGTACGCACCCCTGCCGCCGTTGCCACTACCCGCCACCGGCCACACCCTGACCCG

GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGC CGT CGGGCTCGA  
CATCGGAAGGTGTGGGTGGCGACGGCGCCGGTGGCGGTCTGGACCACGCCGGAGA  
GCGTCGAAGCGGGGGCGGTGTTGCCGAGATCGGCCCGCATGGCGAGTTGAGCGGTTCC  
CGGCTGGCCCGCAGAACAGATGGAAGGCCTCCTGGCGCCGACCGGGCCAAGGGAGCCGC  
GTGGTTCTGGCCACCGTCGGCGTCTGCCGACCAACAGGGCAAGGGTCTGGGAGCGCCG  
TCGTGCTCCCCGGAGTGGAGGCGGGAGCGCGCCGGGTGCCCCGCTCCTGGAGACCTCC  
GCGCCCCGCAACCTCCCCCTACTAGAGCGGCTCGGCTCACCGTCACCGCGACGTCGAGGT  
GCCCGAAGGACCGCGCACCTGGTGATGACCCGCAAGCCC GTGCCTGACGCCGCCACG  
ACCCGCAAGCGCCGACCGAAAGGAGCGCACGACCCATGGCTCCGACCGAAGCCGACCCGGG  
CGGCCCCGCGCACCCGCAACCGCCCCGAGGCCACCGACTCTAGAGTCGGGGCGGCCGG  
CGCTTCGAGCAGACATGATAAGATAACATTGATGAGTTGGACAAACACAACAGAATGCA  
TGAAAAAAATGCTTTATTGTGAAATTGTGATGCTATTGCTTATTGTAACCAATTATAAG  
CTGCAATAAAACAAGTTAACACAACAATTGCAATTCTATTGTTAGGTTAGGGGGAGG  
TGTGGGAGGTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAATCGATAAGGATCAA  
TTCGGCTTCAGGTACCGTCACGATGTTAGGTACGGTCTCGAAGCCGCGTGCGGGTGCCAG  
GGCGTGCCCTTGGCTCCCCGGCGCGTACTCCACCTCACCCATCTGGTCCATCATGATGAA  
CGGGTCGAGGTGGCGGTAGTTGATCCGGCGAACGCGCGCGCACCGGGAAAGCCCTCGCCCT  
CGAAACCGCTGGCGCGGTGGTCACGGTGAGCACGGGACGTGCGACGGCGTCGGGGTGGC  
GATACCGGGGCAGCGTCAGCGGTTCTCGACGGTCACGGCGGGCATGTCGACAGCGAATT  
GATCCGTCACCGATGCCCTTGAGAGCCTCAACCCAGTCAGCTCCTCCGGTGGCGCGGG  
GCATGACTATCGTCGCCGCACTTATGACTGTCTTATCATGCAACTCGTAGGACAGGTG  
CCGGCAGCGCTCTCGCTTCGCTCACTGACTCGCTGCGCTCGGTGTTGGCTGC  
GAGCGGTATCAGCTCACTCAAAGGCGTAATACGGTTATCCACAGAATCAGGGATAACGCA  
GGAAAGAACATG

***Fig. 15***

**pOM IFN-Ins-CMV-pur-attB (SEQ ID NO: 8)**

GGCCGCCACCGCGGTGGAGCTCCAATTGCCCTATAGTGAGTCGTATTACAATTCACTGGCC  
GTCGTTTACACGTCGTGACTGGGAAAACCTGGCGTACCCAACCTTAATGCCCTGCAGC  
ACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCACCGATGCCCTCCAAC  
AGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGCGCATTAGCGGGCGGGT  
GTGGTGGTTACGCGCAGCGTACCGCTACACTTGCCAGCGCCCTAGCGCCGCTCCTTCGC  
TTCTTCCCTTCTCGCCACGTTGCCGGCTTCCCCGTCAAGCTCTAAATCGGGGGC  
TCCCTTAGGGTCCGATTTAGTGTCTTACGGCACCTCGACCCCCAAAAACTGATTAGGGT  
GATGGTTACGTAGTGGCCATGCCCTGATAGACGGTTTTCGCCCTTGACGTTGGAGTC  
CACGTTCTTAATAGTGGACTCTTGTCCAAACTGGAACAAACTCAACCCATTCTCGGTCT  
ATTCTTGATTATAAGGGATTTGCCGATTCGCCATTGGTAAAAATGAGCTGATT  
TAACAAAATTTAACCGCAATTAAACAAAATTTAACGCTTACAATTAGGTGGCACTTT  
CGGGGAAATGTGCGCGAACCCCTATTGTTATTCTAAATACATTCAAATATGTATCC  
GCTCATGAGACAATAACCCTGATAAATGCTCAATAATATTGAAAAAGGAAGAGTATGAGTA  
TTCAACATTCCTGTCGCCCTTATTCCCTTTTGCGGCATTTGCCCTCTGTTTGCT  
CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCAGAGTGGGTTA  
CATCGAACTGGATCTAACAGCGGTAAGATCCTTGAGAGTTTCGCCCGAAGAACGTTTC  
CAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGTATTATCCGTATTGACGCCGGG  
CAAGAGCAACTCGGTGCCGCATACACTATTCTCAGAATGACTGGTTGAGTACTCACCAGT  
CACAGAAAAGCATTTACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCCATAACCA  
TGAGTGATAACACTGCGGCCAACTTACTCTGACAAACGATCGGAGGACCGAAGGGAGCTAAC  
GCTTTTGCAAAACATGGGGATCATGTAACCTGCCCTGATCGTTGGAAACCGGAGCTGAA  
TGAAGCCATACCAAAACGACGAGCGTACACACGATGCCCTGTAAGCAATGGCAACACGTTGC  
GCAAACATTTAACTGGCGAACTACTACTCTAGCTCCCGGAAACAATTAAATAGACTGGATG  
GAGGCGGATAAAAGTTGCAAGGACCACTCTCGCCTGCCCTTCCGGCTGGCTGGTTATTGC  
TGATAAAATCTGGAGGCCGGTGAGCGTGGGTCTCGCGGTATCTTGAGCACTGGGCCAGATG  
GTAAGCCCTCCCGTATCGTAGTTACTACACGACGGGAGTCAGGCAACTATGGATGAACGA  
AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAAGACAAAGT  
TTACTCATATACTTTAGATTGATTAAACTTCAACGTTAAACTGAGTTTCGTTCACTGAGCG  
AGATCCCTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTCGTTCACTGAGCG  
TCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTCTCGCGGTAAAGAGCTAC  
CTGCTGCAAACAAAAAAACACCGCTACACGCGTGGGTTGTTGCCGGATCAAGAGCTAC  
CAACTCTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATAACAAATACTGCTTCTA  
GTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCT  
GCTAATCCTGTTACCACTGGCTGCGTAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT  
CAAGACGATAGTTACCGGATAAGCGCAGCGTGGGCTGAAACGGGGTTCGTCACACAG  
CCCAGCTGGAGCGAACGACCTACACGAACTGAGATACTACAGCGTGAGCTATGAGAAAG  
CGCCACGCTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG  
GAGAGCGCACGAGGGAGCTCCAGGGGAAACGCCCTGGTATCTTATAGTCCTGCGGTT  
CGCCACCTCTGACTTGAGCGTCACTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAA  
AAACGCCAGCAACGCGGCCCTTTACGGTCTGGCTTTGCTGGCTTTGCTCACATGT  
TCTTCCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTGAGTGAGCTGAT  
ACCGCTGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG  
CCCAATACGCAAACCGCTCTCCCGCGTGGCGATTCAATGCACTGGCACGACA  
GGTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTATGTGAGTTAGCTCACTCAT  
TAGGCACCCAGGCTTACACTTATGCTTCCGGCTCGTATGTTGTGAGAATTGTGAGCGG  
ATAACAATTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAAACCTC  
ACTAAAGGGAAACAAAGCTGGGTACCGGGCCCCCTCGACTAGAGGGACAGCCCCCCCCCA  
AAGCCCCCAGGGATGTAATTACGTCCTCCCCGCTAGGGGGCAGCAGCAGGCCGCCGGG  
CTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGGCCAGCGTGCAGGGGACAGCCCG  
GCACGGGGAAAGGTGGCACGGATCGCTTCTGAACGCTTCTCGCTGCTTTGAGCCTG  
CAGACACCTGGGGGATACTGGGGAAAAGCTTACGGCTGAAAGAGAGATTAGAATGACAGA  
ATCATAGAACGGCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTAT  
GTGCAAGGGTCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG

CCTGCAGGGATGGGGCATCCACAGCCTGGCAACCTGTTAGTCAGTCGTACCCACCCTCT  
GGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCAGTGTAAAGCCATTC  
CCCCTTGTCTATCAAGGGGGAGTTGCTGACATTGTTGGCTGGGGGACACATGTTG  
CCAATTCACTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGAC  
GTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCAGTCAGAACAGCC  
TTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAG  
CACAAAAAGGCCACAGACACTGCTGGTCCCTGTCAGCCTGCATGTTGATGGTGTCTG  
GATGCAAGCAGAAGGGTGGAAAGAGCTTGCGAGAGATAACAGCTGGGTAGTAGGACTGG  
GACAGGGAGCTGGAGAATTGCATGTAGATGTTCATACAATCGTCAAATCATGAAGGCTGGA  
AAAGCCCTCAAGATCCCCAAGACCAACCCACCCACCGTGCCTACTGCCATGTCC  
CTCAGTGCACATCCCCACAGTTCTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT  
GGCAGCTGTGCCACTGCAGCACCGCTTTGGAGAAGGTAATCTGCTAAATCCAGCCCG  
ACCCCTCCCTGGCACACGTAAAGGCCATTATCTCATCCAACCTCCAGGACGGAGTCAGTGA  
GGATGGGGCTCTAGTCGAGGTCGACGGTATCGATAAGCTGATTAGGCAGAGCAATAGGACT  
CTCAACCTCGTGAGTATGGCAGCATGTTAATCTGCACTGGAGTCCAGCGTGGAAACAATC  
TGCCTTGACATGAGTCTCGTGGCCAATATTCCCCAACGGTTTCTCAGCTGTCTG  
TCTCCTAAGCTCTCAAAACACCTTTGGTGAATAAAACTCACTGGCAACGTTATCTGTCT  
TACCTTAGTGTACGTTCATCCCTATTCCCTTCTCCTCCGTGTGGTACACAGTGGT  
GCACACTGGTTCTCTGTTGATGTTCTGCTGACAGCCAATGTGGGAAAGTTCTTCCTGC  
CACGTGTCTGTGTTGTTCACTCAAAAAGGGCCCTGGCTCCCTGGAGCTCTCAGGCA  
TTTCCTTAATCATCACAGTCAGCTGGCAGGATTAGTCCCTCTAAACCTTAGAATGACCTG  
AACGTGTGCTCCCTTTGTAGTCAGTGCAGGGAGACGTTGCCTCAAGATCAGGGTCCATC  
TCACCCACAGGGCATTCCAAGATGAGGTGGATGGTTACTCTCACAAAAAGTTTCTTAT  
GTTGGCTAGAAAGGAGAACTCACTGCCTACCTGTGAATTCCCTAGTCCTGGTTCTGCTGC  
CACTGCTGCCTGTGCAGCCTGTCAGGCTGCCCAGGATTAGTCCCTCTAAACCTTAGAATGACCTG  
CACCCCTGTCTCCACTGAAATGACCTCAGTGCACGTGTTGTATAGGGTATAAGTACGGGAG  
GGGGATGCCGGCTCCCTCAGGGTTGCAGAGCAGAAGTGTCTGTATAGAGTGTGTTA  
ATCTATTAATGTAACAGAACAACTTCAGTCCTAGTGTGTTGTTGCTGGAATTGCCATGTG  
GTAGGGACAGGCCTGCTAAATCACTGCAATGCCTATGTTCTGAAGGTATTGGGAAAGAAA  
GGGATTGGGGATTGCCGTGTGATTGGCTTAAATTGAATGGCAAATCACAGGAAAGCAGTTC  
TGCTCAACAGTTGGTTCTCAGCCAATTCTGCAGCCAAAGAGCCGGGTGCCAGCGATAT  
AATAGTTGTCACTTGTGCTGTATGGATGACAGGGAGGTAGGGTGACCTGAGGACCACCC  
CAGCTCTGCTAGCGTAGGTACAGTCACCACCTCCAGCTCCACACAGGTCCATCGTGGTT  
ACCAAAGAAACACAATTATTGGACCAGTTGGAAAGTCACCCGCTGAATTGTGAGGCTAGA  
TTAATAGAGCTGAAGAGCAAATGTCCTCAACTTGGAGATACTAGTTGGTATTAGTATCAGAG  
GAACAGGGCCATAGCACCCTCATGCTATTAGATTGGCTGGCATGTACTTCAAGATGAT  
TTGTAACTAACATGGCTTATTGTCCTGCTTAAGTCTGTGCTCTTCAAGTAAATGTTCTT  
TGGTTATATAACCTTCTGCCATTGCTCTTCAGGGTTCTGCAGAACACTGGCTGCTT  
AATCTAGTTAACTGTTGCTGATTCTTAGGGATAAGATCTGAATAAAACTTTTGTGGC  
TTTGGCAGACTTCTAGCTGGCTTAGCTCCACATTAGCTTTGCTGCCTTCTGTGAAGC  
TATCAAGATCCTACTCAATGACATTAGCTGGGTGCAGGTGTACCAAATCCTGCTCTGTGGAA  
CACATTGTCATGATGACACCGAAGGCAAACGTGAACCTCAAAGAGGCACAGAGTTAAGAAGAAG  
TCTGTCAATTAGAGGAAAGCCAAAGTGGCATTAGACACACTTTCCATGCAGCAGCATTC  
CAGTAGGTTCATATAAAACTACAAAATGGAATAAACCACTACAAATGGGAAAGCCTGATA  
CTAGAATTAAATATTCAACCCAGGCTCAAGGGGTGTTCATGGAGTAATATCACTCTATAAA  
AGTAGGGCAGCCAATTATTCAACAGACAAAGCTTTTTCTGTGCTGCAGTGTGTTTT  
CGGCTGATCCAGGGTACTTATTGTGGGTCTGAGAGCTGAATGATTCTCCTTGTGCTG  
TGGTGAAGGAGATATGCCAGGGGAGATGAGCATGTTCAAGAGGAAACGGTGCATTGGT  
GGCTTGGGAGAAAGGTAGAACGATATCAGGTCCATAGTGTCAACTAAGAGATCTGAAGGATGG  
TTTACAGAACAGTTGACTGGCTGGGTGCAGGCTGGCTGTAATGGATGGAAGGATGGAC  
AGATGGGTGGACAGAGATTCTGTGCAGGGAGATCATCTCCTGAGCTGGTGCCTGACAGACT  
GCAGATCCATCCCATAACCTCTCCAGCATGAGAGCGCGGGGAGCTTGGTACTGTTCAGTC  
TGCTGCTTGTGCTTCCCTGGGTGCACAGTGGTGAATTCTTACTCACACAGGGAAAAACCT  
GAGCAGCTCAAAGTGAACAGGTTGCTCATAGGCCATTCAAGTGTCAAGATGAGGTTTT

GGTTCTGTTTGTAAAGGTGGAGAAGCACTGAAGGATCAGTTGCAGGGCAGGGTTA  
GCACTGTTAGAGAAGTCTTATTAACTCCTCATGAACAAAAAGAGATGCAGGTGCAGA  
TTCTGGCAAGCATGCAGTGAAGGAGAAAGCCCTGAATTCTGATATATGTGCAATGTTGGC  
ACCTAACATTCCCCGCTGAAGCACAGCAGCTCCAGCTCATGCAGTACTCACAGCTGGTGCA  
GCCCTCGGCTCCAGGGTCTGAGCAGTGCAGGACTCACAGGTTCCATGTCTTACACTGA  
TAATGGTCCAATTCTGGAATGGGTGCCATCCTGGAGGTCCCAGGCCAGGCTGGCTGC  
GTCTCCGAGCAGCCCAGTGGTGGTAGCCAGCCATGGCAGGAGTTAGACCTGAGTG  
GTCTTAAGGTCCCTCAACCTAACGCCATCCTACGATTCTAGGAATCATGACTTGTGAGTG  
TGTATTGCAGAGGAATAATTAAAGTTATAATGTTCTCCCTCCTGTTGTCAAAG  
TTATCTGATGCCATTATCAATGCTTTGGAGTCTCAGTCATTCTTCTTACAMCAAAAAGA  
GGAGGAAGAATGAAGAGAATCATTAAATTCTGATGAATAGTAGGATTAGAAAGCTGTA  
CGTAATGCCGTCCTTGTATCGAGCTGTAAGGTTCTCATCTTACAGCGTGGTACATA  
TCAGCACTTCCATCTGATGTGGAAAAAAAAACCTTATCATCTACAGTCTCTGTACCTAA  
ACATCGCTCAGACTCTTACCAAAAAAGCTATAGGTTAAAACATACATCTGCTGATAATT  
GCCTGTTTAGCTCTTCCATATGCTGCGTTGTGAGAGGTGCGTGGATGGGCTAAAC  
TCTCAGCTGCTGAGCTGATGGGTGCTTAAGAATGAAGCACTCACTGCTGAAACTGTTTCA  
TTTCACAGGAATGTTTAGTGGATTGTTATAACTACATATTCCCTCAGATAAAATGAAAT  
CCAGAAATAATTATGCAAACACTGCATCCGTTGCACAGGCTTTATCTGCTAGCAAAGGA  
AATAATTGGGGATGGCAAAACATTCTCAGACATCTATATTAAAGGAATATAATCCTG  
GTACCCACCCACTTCATCCCTCATTATGTTACACTCAGAGATACTCATTCTTGTGTTA  
TCATTGATAGCGTTCTTGGTTCTTGCCACGCTCTGGCTATGGCTGCACGCTCTGCA  
CTGATCAGCAAGTAGATGCGAGGGAAAGCAGCAGTGCAGGAGGGCTGCCCTCAGCTGGCACCCA  
GCCGCTCAGCCTAGGAGGGGACCTGCCTTCCACCAGCTGAGGTGCGAGCCCTACAAGCTTA  
CACGTGCTGCGAGCAGGTGAGCAAAGGAGTCTTCATGGTGTGTTCTGCTGCCCGGAAGC  
AAAACTTTACTTTCATTCCCTTAAGGCTCAGCCAGGGCCATTGCTGAGGACGGCATCGGG  
CGTTCAATTCTCTTCCCTTAAGGCTCAGCCAGGGCCATTGCTGAGGACGGCATCGGG  
GCCCTGGACCAATCTGTCAGGAGGAGCAGATGGTTCACTACATCAGTGGATGTTGCT  
GCCTGTAATGTCCTCTGAAGGAAGAACGTGCCCTCCAAGTGCCAGCCCCACAGCCCC  
AGCCCCCTCCCTGTGCTGCTCCAATTCTCATCTCCCTCTCCCTCTCCCTTGCTGTTGTC  
TCGGGTAGAAATCATGAAGATTAGAAGAGAAAACAAAATAACTGGAGTGGAAACCCAGGTG  
ATGCAGTTCAATTCAAGCTGTCATAGGTTGCGTTGCTATAGGTCTGTATCAGAGATGCTARC  
ACCACTTGTGCTGCGGTGCTTAACTCGGGTGAACCTCCTTCACTCGCATATTGCGGGCC  
TTATTTACATCCCCAGCATCCATCACCCTCTGGAAATGGGCGACTGGATCTCTAATGGA  
AGACTTCCCTTTCAGAGCCTGTTGAGGATGTGCAAGAAAGTGGAGGGCTGAGC  
AGCAGCACTGCCCTCAGGGAGCAGGAGCGGATGCCATCGGTGGCAGCATCCAAATGATGTC  
AGCGGATGCTGAGCAGGGCAGCGGAGCAGAAGGGACAGAACGGCAGTGCACCTCTGTTGACA  
TGGTATTGGCAGCGATTAAACACTCGCTCTAGTCCTGCTATTCTCACAGGCTGCATTC  
AAATGAACGAAGGGAAAGGGAGGCAAAAGATGCAAAATCCGAGACAAGCAGCAGAAATATT  
CTTCGCTACGGAAGCGTGCACAAACACCTCTCCAAACAGCACAGCACAGCGTAAC  
CTTTTCAAGACCAGAAAAGGAAATTCAAAAGCCTCTGTGGATACCAGCGCTCAGCTCT  
CCTGATAGCAGATTCTTGTCAAGGTTGCAATGGGTATGGTGCAGGAGGTGAGGGACCA  
TATGATCATATACAGCACAGCAGTCATTGTCATGTATTAAATATATTGAGTAGCAGTGT  
ACTTTGCCAAAGCAATAGTTAGAGATGAGTCCTGCTGCATACCTCTATCTTAAACTA  
TATAAAATAGTAAACCTCTCAGTTCAAGCCACGTGCTCTCTGTGTCAGCACCAATGGTCT  
TCGCCTGCACCCAGCTGCAAGGAATCAGCCCGTGAATCTCATTAAACACTCAGCTCTGCAGGAT  
AAATTAGATTGTCAGCTACAGAAAAGTCTCCACTCTTGTGTTAATTACAGCGGAACAATTGTT  
CAGTGCAGTCTCTGCTGCTTCTGCGCCAGCAAACAGTGC  
GGTCTAATTGTCAGCTACAGAAAAGTCTCCACTCTTGTGTTAATTACAGCGGAACAATTGTT  
CAGTGCAGTCTCTGCTGCTTCTGCGCCAGCAAACAGTGC  
CAGGCTATAGCAGCGTGAATGCTACCTCTCACTCCATCCTCTCTCTCTCTCTCTCTCT  
GAGAGCTGTTGTTCACTCTCAGCCACTCTGAACAATACCAAAACTGCTACGCACTGCCTC  
CCTCGGAAAGAGAATCCCTTGTGCTTTTATTACAGGATCCTCTTAAAGCAGACC  
ATCATTCACTGCAAACCCAGAGCTTCACTGCCTCTCCACAAACGAAAACAGCCGGCTTC  
ATTGTCCTTTAAATGCTGTTCCAGGTGAATTGGCCAGCGTGTGGCTGAGATCCA  
GGAGCACGTGTCAGCTTCTGCTCTCATTGCTCCTGCTGATTGCTCTTCTGGGGTT  
CCAAGAGGGGGGAGACTTGCGCGGGATGAGATAATGCCCTTCTAGGGTGGCTGCT

GGGCAGCAGAGTGGCTGGGTCAGTGACCAATGGGAGGCACCAGTGGGGGTGTT  
TGTGCAGGGGGAAAGCATTACAGAATGGGCTGATCCTGAAGCTGAGTCCAAGGCTTG  
TCTGTGTACCCAGTGAAATCCTCCTGTTACATAAAGCCCAGATAGGACTCAGAAATGTA  
GTCATTCCAGCCCCCTCTCCTCAGATCTGGAGCAGCACTTGTGAGGCCAGTCCTCCCC  
AAAATGCACAGACACCTCGCCGAGTGGAGGGAGATGTAACACAGCGAAGGTTAATTACCTCCTG  
TCAAAAAACACTTTGTGGTCCATAGATGTTCTGTCATCTTACAAAACAGAACCGAGAGGCA  
GCGAGCACTGAAGAGCGTGTCCCAGTGTGAGTTAATGAGACTTGGCAGCTCGCTGTGAGA  
GATGATCCCTGTGCTTCATGGAGGCTGTAACCTGTCTCCCCATCGCCTTCACACCGCAGTG  
CTGTCCTGGACACCTCACCCCTCATAAGCTGTAGGATGCAGCTGCCAGGGATCAAGAGACT  
TTTCCTAAGGCTCTAGGACTCATTTGCCGCTCAGTAGCGTGCAGCAATTACTCATCCCA  
ACTATACTGAATGGGTTCTGCCAGCTCTGCTTGTCAATAAGCATTCTCATTGCA  
CTCTAAGTTCTCAGCAGCACCGCTCTGGGTGACCTGAGTGGCCACCTGGAAACCGAGGG  
GCACAGCCACCACTCCCTGTTGCTGCTCCAGGGACTCATGTGCTGCTGGATGGGGGA  
AGCATGAAGTTCTCACCCAGACACCTGGGTTGCAATGGCTGCAGCGTGCCTTCTGGTAT  
GCAGATTGTTCCAGCCATTACTGTAGAAATGTGCTGTTGAAAGGCCCTTGATCTCTTCT  
GTGGCCCTTCAGCAAAAGCTGTGGGAAAGCTCTGAGGCTGTTCTGGGTGTTGGAGGAAT  
TGTATGTTCTCTTTAACAAAATTATCCTTAGGAGAGAGCACTGTGCAAGCATTGTCAC  
ATAAAACAATTCAAGGTTGAAAGGGCTCTGGAGGTTTCCAGCCTGACTACTGCTCGAAGCA  
AGGCCAGGTTCAAAGATGGCTCAGGATGCTGTGCTTCCCTGATTATCTGTGCCACCAATG  
GAGGAGATTCACAGCCACTCTGCTCCGTGCCACTCATGGAGAGGAATATTCCCTTATATT  
CAGATAGAATGTTATCCTTAGCTCAGCCTCCCTATAACCCATGAGGGAGCTGCAGATCC  
CCATACTCTCCCTCTGGGGTGAAGGCCGTGCCCCAGCCCCCTCCACCCCTGTG  
CCTAAGCAGCCGCTGGCTCTGCTGGATGTTGCTTATATGTCATGCCTGCTTGCAGT  
CCAGCCTGGACATTAAATTACACCAGGGTAATGTGAACTGTGTCATCTCCCTGCAG  
GGTACAAAGTTCTGCACGGGTCTTCGGTTCAAGAAAACCTTCACTGGTGCTACCTGAAT  
CAAGCTCTATTAAATAAGTTATAAGCACATGGATGTTCTAGAGATACTGTTAAATG  
GTATCAGTGAATTATTGCTTGTGCTTACTTCAAACAGTGCCTTGGCAGGAGGTGA  
GGGACGGGTCTGCCGTTGGCTCTGCACTGATTCTCCAGGCGTGTGGCTCAGGTCAAGATA  
GGTCACTCTGTGGCCAGAAGAAGGACAAAGATGGAATTGCAAGATTGAGTCACGTTAACGAG  
GCATCTGGAGTGATTGAGGCAGTTCATGAAAGAGCTACGACCAACTTATTGTTGTTTCC  
CCTTTACAACAGAAGTTTACAAACAGTGGCAAAGGCCAGGAATGTTGGAAAAG  
TGTAGTTAAATGTTGTAATTCAATTGCTGGAGTGCTACCGACTAAGAAAAAGTCTACC  
TTTGGTATGGTAGTCCTGCAGAGAAATACACATCAATATTAGTTGGAAAAAACACCA  
CCACCAAGAAACTGTAATGGAAATGTAACAGAAGAAATTCTGGTAAGAGAGAAAGGATG  
TCGTATACTGGCCAAGTCCCTGCCAGCTGTCAGCCTGCTGACCCCTGCACTTCAGGACCAT  
GAAACGTGGCACTGTAAGACGTGCTCCCTGCTTGTGCTGCCACAGATCTGCCCCCTGTG  
CTGACTCCTGCACACAAGAGCATTCCCTGTAGCCAAACAGCGATTAGCCATAAGCTGCACC  
TGACTTTGAGGATTAAGAGATTGCAATTAAAGTGGATTGCAAGCAGGAGATCAGTGGCAGGGTT  
GCAGATGAAATCTTTCTAGGGTAGCTAAGGGCTGAGCAACCTGTCCTACAGCACAAAGCC  
AAACCAAGCCAAGGGTTTCTGTGCTGTTACAGAGGCAGGGCCAGCTGGAGCTGGAGGAGG  
TTGTGCTGGGACCCCTCTCCCTGTGCTGAGAATGGAGTGATTCTGGGTGCTGTTCTGTG  
CTTGCACTGAGCAGCTCAAGGGAGATCGGTGCTCCTCATGCAGTGCCAAACTCGTGTG  
TGCAGAAAAGATGGATGTCACCTCCCTGCTAATGCAGCCGTGAGCTTATGAAGGCAATG  
AGCCCTCAGTGCAGCAGGAGCTGACTGCACTCTGTAGGTGCTAGGGAAAATCTGTG  
CCAGGGATGCATTCAAGGGCAATATATCTGAGGCTGCCAAATCTTCTGAAATATTC  
ATGCGTGTCCCTTAATTATAGAAACAAACACAGCAGAATAATTATTCACAGCCTCC  
CGAAGGAAACCCATATTCCATGTAGAAATGTAACCTATATAACACAGCCATGCTGCATCC  
TTCAGAACGTGCCAGTGCCTCATCTCCATGGCAAAACTACAGGTATTCTCACTATGTTG  
ACCTGTGAAAGGAACCATGGTAAGAAACTTCGGTTAAAGGTATGGCTGCAAAACTACTCATA  
CCAAAACAGCAGAGCTCCAGACCTCCCTTAGGAAAGAGGCCACTGGAGAGGGATGGTGTGA  
AGGCTGGAGGTGAGAGACAGAGCCTGCTCCAGTTCTGTCTCTATTCTGAAACGTTG  
CAGGAGGAAAGGACAACACTGTACTTTCAGGCATAGCTGGTGCCTCACGTAATAAGTCCCC  
GAACTCTGTGTCATTGTTCTTAAGATGCTTGGCAGAACACTTGTAGTCATTGCTAA  
CTGTGACTAGGTCTGTAATAAGTGCCTCGTGTAAAGGTTCAAGTGAACATTAGTGG

TATTTGACAGCATTACCTTGCTTCAAGTCTTCTACCAAGCTCTTCTATACTTAAGCAGTG  
AAACCGCCAAGAAAACCCTCCTTTATCAAGCTAGTGCTAAATACCATTAACCTCATAGGTT  
AGATACGGTGCTGCCAGCTCACCTGGCAGTGGTTGGTCAGTTCTGCTGGTGACAAAGCCTC  
CCTGGCCTGTGCTTTACCTAGAGGTGAATATCCAAGAATGCAGAACTGCATGGAAAGCAGA  
GCTGCAGGACAGATGGTGCTGAGCCTTAGCTGCTCCTGCTGGGAGATGTGGATGCAGAGAC  
GAATGAAGGACCTGTCCCCCTACTCCCCCTCAGCATTCTGTGCTATTAGGGTTCTACCAAGAGT  
CCTTAAGAGGTTTTTTTTGGTCCAAAAGTCTGTTGGTTTGACCACTGAGA  
GCATGTGACACTTGTCTCAAGCTATTACCAAGTGTCCAGCAGGAAATCAATTGCCTGGGAGA  
CGCAGACCATTACCTGGAGGTCAAGGACCTCAATAAATATTACCAAGCCTCATTGTGCCGCTGA  
CAGATTCAAGCTGGCTGCTCCGTGTTCCAGTCCAACAGTCGGACGCCAGTTGTATATATT  
TGCAGGCAGCCTGGGGGACCATCTCAGGAGCAGAGCAGCCGAGCCGCTGCAGAGCCGG  
GCAGTACCTCACCATGGCTTGACCTTGCCTACTGGTGGCTCTCCTGGTGCTGAGCTGCA  
AGAGCAGCTGCTGTGGCTGCCATCTGCCTCAGACCCACAGCCTGGGAGCAGGAGGACC  
CTGATGCTGCTGGCTCAGATGAGGAGAATCAGCCTGTTAGCTGCCTGAAGGATAGGCACGA  
TTTGCTTCCCAAGAGGAGTTGGCAACCAGTTAGAAGGCTGAGGACATCCCTGTGC  
TGCACCGAGATGATCCAGCAGATCTTAACCTGTTAGCAGCAAGGATAGCAGCGCTGCTTGG  
GATGAGACCTGCTGGATAAGTTTACACCGAGCTGTACCAAGCAGCTGAACGATCTGGAGGC  
TTGCGTGATCCAGGGCGTGGCGTGACCGAGACCCCTCTGATGAAGGAGGATAGCATTGCG  
CTGTGAGGAAGTACTTCAGAGGATCACCTGTACCTGAAGGAGAAGAAGTACAGCCCCTGC  
GCTTGGGAAGTCGTGAGGGCTGAGATCATGAGGAGCTTAGCCTGAGCACCAACCTGCAAGA  
GAGCTTGAGGTCTAAGGAGTAAAAGTCTAGAGTCGGGGCGGCCGCTCGAGCAGACA  
TGATAAGATACTTGATGAGTTGGACAAACCACAATAGAACATGCACTGAAAAAAATGCTT  
ATTGTGAAATTGTGATGCTATTGCTTATTGTAACCATTATAAGCTGCAATAAACAAAGT  
TAACAACAACAATTGCATTCTATTATGTTAGGTTAGGCAATTAGCCATATTAGTCATTG  
AAAGCAAGTAAAACCTCTACAAATGTGGTAAATCGATACCGTCGACCTCGACTAGAGCGGC  
CACTAACATACTGCTCTCCATCAAACAAAACAAACTAGCAAAATAGGCTGTC  
CCCAGTGCAAGTGCAGGTGCCAGAACATTCTCATCGATAGGTACCGAGCTCTACGCGTG  
CTAGCCCTCGAGCAGGATCTACATTGAAATCAATTGGCAATTAGCCATATTAGTCATTG  
GTTATATAGCATAATCAATATTGGCTATTGGCCATTGCACTACGGTATCTATATCATAAT  
ATGTACATTATATTGGCTCATGTCCAATATGACCCCATGTTGACATTGATTATTGACTAG  
TTATTAATAGTAATCAATTACGGGGTATTAGTCATAGCCCATATATGGAGTTCCGCGTTA  
CATAACTACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCCCGCCATTGACGTCA  
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGTGGA  
GTATTTCACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCC  
CTATTGACGTCAATGGGAGTTGTTGGCACCCTAACACGGACTTCCAAAATGTCGT  
GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGGTT  
TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTCCAAGTCTCCACC  
CCATTGACGTCAATGGGAGTTGTTGGCACCCTAACACGGACTTCCAAAATGTCGT  
AACAAACTCCGCCCCATTGACGCAAATGGGCGTAGGGCGTGTACGGTGGGAGGTCTATAAAG  
CAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTGACCTCC  
ATAGAAGACACCGGGACCGATCCAGCCTCCCTCGAAGCTCGACTCTAGGGCTCGAGATCT  
GCGATCTAAGTAAGCTGCATGCCCTGCAGGTGGCCACGACCGGTGCCACCATCCC  
CTGACCCACGCCCTGACCCCTCACAAAGGAGACGACCTCCATGACCGAGTACAAGCCCACG  
GTGCGCCTGCCACCGCGACGACGTCCCCCGGCCGTACGCACCCCTGCCCGCGTTCGC  
CGACTACCCCGCCACCGCGCACACCGTCGACCCGGACGCCACATCGAGCGGGTCACCGAGC  
TGCAAGAACTCTTCCTCACCGCGTCGGCTCGACATCGCAAGGTGTGGGTCCGGAGGAC  
GGCGCCCGGGTGGCGGTCTGGACCCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTGCCGA  
GATCGGCCCCGCGCATGGCGAGTTGAGCGGTTCCCCGCTGGCCCGCAGCAACAGATGGAAG  
GCCTCTGGCGCCGACCCGGCCAAGGAGCCCGCGTGGTTCTGGCCACCGTCGGCGTCTCG  
CCCGACCACCAAGGGCAAGGGTCTGGCAGCGCCGTCGTGCTCCCGGAGTGGAGGCGGCCGA  
GCGCGCCGGGGTGGCCCGCTTCCCTGGAGACCTCCGCCACGCCAACCTCCCTCTACGAGC  
GGCTGGCTTCACCGTCACCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTGGTGCATG  
ACCCGCAAGCCCGGTGCCGTGACGCCGCCCCACGACCCCGCAGCGCCGACCGAAAGGAGCGC  
ACGACCCCATGGCTCCGACCGAAGCCGACCCGGGGCCCCGCCGACCCCGCACCCGCCCC

GAGGCCACCGACTCTAGAGTCGGGGCGGCCGCTTCGAGCAGACATGATAAGATAACAT  
TGATGAGTTGGACAAACCACAACCTAGAATGCAGTGAAAAAAATGCTTATTGTGAAATT  
GTGATGCTATTGCTTATTGTAACCATTATAAGCTGAATAAACAGTTAACAAACAACAT  
TGCATTCACTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTAAAGCAAGTAAAA  
CCTCTACAAATGTGGTAAAATCGATAAGGATCAATTGGCTTCAGGTACCGTCGACGATGTA  
GGTCACGGTCTCGAAGCCGGTGGGTGCCAGGGCGTGCCTTGGGCTCCCCGGCGCGT  
ACTCCACCTCACCATCTGGTCCATCATGATGAACGGGTGAGGTGGCGGTAGTTGATCCCG  
GCGAACGCGCGCGCACCGGGAAAGCCCTGCCCTCGAAACCGCTGGGCGCGTGGTCACGGT  
GAGCACGGGACGTGCGACGGCGTCGGCGGTGCGGATAACGCGGGCAGCGTCAGCGGGTTCT  
CGACGGTCACGGCGGGCATGTCGACAGCGAATTGATCCGTCGACCGATGCCCTGAGAGCC  
TTCAACCCAGTCAGCTCCCTCCGGTGGCGGGGATGACTATCGTCGCCGCACTTATGAC  
TGTCTTCTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC

*Fig. 16*

**pRSV-C31int (SEQ ID NO: 9)**

CTGCATTAATGAATCGGCCAACGCGCGGGAGAGCGGTTTGCCTTCGCTATTGGCGCTCTTCC  
GCTTCCTCGCTCACTGACTCGCTCGCTCGGTCTCGCTCGGCTCGGCAGCGGTATCAGCT  
CACTCAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGATAACGCAGGAAAGAACATG  
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAAGGCCGCTGCTGGCGTTTTTC  
CATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGA  
AACCCGACAGGACTATAAAGATAACCAGCGTTCCCGCTTGGAAAGCTCCCTCGCTCGCTCT  
CCTGTTCCGACCTGCCGCTTACCGATACTGTCCGCCTTCTCCCTCGGGAAAGCGTG  
GCGCTTCTCAATGCTACGCTGTAGGTATCTCAGTTGGTAGGTCTGCTCCAG  
CTGGGCTGTGTCACGAACCCCCGTTCAGGCCGACCGCTGCCCTTATCCGTAAC  
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAAC  
AGGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTGTAAGTGGTGGCCTAAC  
TACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGGCCAGTTACCTTC  
GGAAAAAGAGTTGGTAGCTTGTGATCCGCAAACAAACACCAGCTGGTAGCGGTGGTTTT  
TTTGGTTGCAAGCAGCAGATTACCGCAGAAAAAAAGGATCTCAAGAAAGATCCTTGATC  
TTTCTACGGGCTCTGACGCTCAGTGGAACGAAAACACGTTAACGGATTGGTCATG  
AGATTATCAAAAGGATCTCACCTAGATCCTTAAATTAAATGAAGTTAAATCA  
ATCTAAAGTATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA  
CCTATCTACGGATCTGCTATTGTTCATCCATAGTGCCTGACTCCCGTCTGTAG  
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCGAGTGCTGCAATGATAACCGCAGAC  
CCACGCTCACCGGCTCCAGATTATCAGCAATAAACCAGCCAGCCGGAAAGGCCAGCGC  
AGAAGTGGTCTGCAACTTATCCGCTCCATCCAGTCTATTAAATTGTTGCCGGAAAGCT  
AGAGTAAGTAGTCGCCAGTTAATAGTTGCGAACGTTGGCCATTGCTACAGGCATC  
GTGGTGTACGCTCGTCTGGTATGGCTTCACTGAGCTCCGGTCCAAACGATCAAGG  
CGAGTTACATGATCCCCATGTTGCAAAAAAGCGGTTAGCTCCTCGTCTCCGATC  
GTTGTCAGAAGTAAGTGGCCGAGTGTATCACTCATGGTTATGGCAGCACTGCATAAT  
TCTCTACTGTCATGCCATCCGTAAGATGTTCTGTACTGGTGAGTACTCAACCAAG  
TCATTCTGAGAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGAT  
AATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCATATTGGAAAACGTTCTCGGGG  
CGAAAACCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGAGTAAACCACTCGTGC  
CCCAACTGATCTCAGCATCTTACTTACCAAGCGTTCTGGGTGAGCAAAACAGGA  
AGGCAAAATGCCGCAAAAAGGGAATAAGGGCGACACGAAATGTTGAATACTCATACTC  
TTCCCTTTCAATTATTGAGCATTATCAGGGTTATTGTCATGAGCGGATACATA  
TTTGAATGTATTAGAAAATAACAAATAGGGTCCCGCACATTCCCCGAAAAGTG  
CCACCTGACGTCGACGGATCGGGAGATCTCCGATCCCCATGGTCACTCTCAGTACAA  
TCTGCTCTGATGCCGCTAGTTAACCGAGTATCTGCTCCCTGCTGTGTTGGAGGTCG  
CTGAGTAGTGCGCGAGCAAAATTAAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCA  
TGAAGAATCTGCTTAGGGTTAGGCCTTTCGCTGCTCGCGATGTACGGGCCAGATATA  
CGCGTCTAGGGTCTAGGATCGATTCTAGGAATTCTCTAGCCGGTCTAGGGATCCCG  
GCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAACGCACT  
GCTCCCTGCTTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTAAAGCTACAAC  
AAGGCAAGGCTTGACCGACAATTGCAAGGAGAGAAAAGCACCCTGCATGCCGATTGGTG  
GCTTCGCGATGTACGGGCCAGATATACCGCTATCTGAGGGACTAGGGTGTGTTAGGCG  
AAAAGCGGGCTTCGGTTACCGGTTAGGAGTCCCTCAGGATATAGTAGTTCGCTT  
TTGCATAGGGAGGGAAATGTAGTCTTATGCAATACACTGTAGTCTGCAACATGGTA  
ACGATGAGTTAGCAACATGCCCTACAAGGAGAGAAAAGCACCCTGCATGCCGATTGGTG  
GAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGCAACAGACAGGTCTGACATGGATT  
GGACGAACCACTGAATTCCGCATTGACGAGATAATTGTATTAAAGTGCTAGCTCGATAC  
AATAAACGCCATTGACCATTACACATGGTGCACCTCAAGCTTGCATGCCGCT  
GGTACCGGTCCGGAATTCCGGGTGACGAGCTCACTAGTCAGTCTGAGGGTCTGCCGACATGAC  
ACAAGGGGTTGTGACCGGGTGGACACGTACCGGGTCTACGACCGTCAGTCGCGCGA  
GCGCGAGAATTGAGCGCAGCAAGCCCAGCGACACAGCGTAGCGCCAACGAAGACAAGGC  
GGCCGACCTTCAGCGCGAAGTCGAGCGCAGCGGGGCCGGTTCAAGGTTCTGCGGCCATTT  
CAGCGAAGCGCCGGCACGTCGGCTCGGGACGGCGAGCGCCGGAGTTCGAACGCAT

CCTGAACGAATGCCGCCGGCGGCTAACATGATCATTGTCTATGACGTGTCGCCTT  
CTCGCGCTGAAGGTCAATGGACCGATTCCGATTGTCTCGAACCGCTGGCGT  
GACGATTGTTCCACTCAGGAAGCGTCTCCGGCAGGGAAACGTCAATGGACCTGATTCA  
CCTGATTATGCGGCTCGACCGTCGCACAAAGAATCTTCGCTGAAGTCGGCGAAGATTCT  
CGACACGAAGAACCTTCAGCGAAATTGGCGGGTACGTGGCGGGAAAGGCGCCTTACGG  
CTTCGAGCTTGGTGGTGGCGTGGAGATCAAGACGACAAACACCTTCCCTCAAGCCGG  
CATCAACAAGCTTGCACACTCGACCAACTCCCCTTACCGGACCCCTCGAGTTCGAGGCCGA  
CGTAATCCGGTGGTGGTGGCGTGGAGATCAAGACGACAAACACCTTCCCTCAAGCCGG  
CAGTCAAGCCGCATTACCCGGCAGCATTACGGGCTTGTAAGCGCATGGACGCTGA  
CGCCGTGCCGACCCGGGGCGAGACGATTGGGAAGAACCGCTTCAAGCGCCTGGGACCC  
GGCAACCGTTATGCAATCCTTCGGGACCCCGTATTGCGGGCTTCGCGCTGAGGTGAT  
CTACAAGAAGAACCGGACGGCACGCCACCAGAAGATTGAGGGTTACCGCATTAGCG  
CGACCCGATCACGCTCCGGCCGGTCAAGCTTGATTGCGGACCGATCATCGAGCCGCTGA  
GTGGTATGAGCTTCAGGCCTGGTGGACGGCAGGGGGCGCGCAAGGGGCTTCCCGGGG  
GCAAGCCATTCTGTCCGCCATGGACAAGCTGTACTCGCAGTGTGGCGCGTCACTGACTTC  
GAAGCGCGGGAAAGAACGATCAAGGACTCTTACCGCTGCCGTGCCCGAAGGTGGTCGA  
CCCGTCCGCACCTGGCAGCACGAAGGCACGTCAACGTCAAGCATGGCGGACTCGACAA  
GTTCGTTGCCAACGCATCTCAACAAGATCAGGCACGCCAACGGCGACGAAGAGACGTT  
GGCGCTCTGTGGGAAGCCGCCGACGCTTCCGCAAGCTCACTGAGGCCCTGAGAAGAG  
CGCGAACGGCGAACCTTGGTGGGAGCGCGCCGACGCCCTGAACGCCCTGAAGAG  
GTACGAAGACCGCGCGCAGGCGCTACGACGGACCCGTTGGCAGGAAGACACTTCCGAA  
GCAACAGGCAGCGCTGACGCTCCGGCAGCAAGGGCGGAAGAGCGGCTTGCCTGA  
AGCCGCCGAAGCCCCGAAGCTTCCCTGACCAATGGTCCCCGAAGACGCCGACGCTGA  
CCCGACGGCCCTAACGCTGGTGGGGCGCGCTCAGTAGACGACAAGCGCTGGT  
CGGGCTCTCGTAGACAAGATCGTTGTCACGAAGTCGACTACGGCAGGGGGCAGGGAA  
GCCCATCGAGAACGCGCTTCGATCACGTGGCGAACGCCGCCGACGACGACGAAGA  
CGACGCCAGGACGGCACGGAAGACGTAGCGCGTAGCGAGACACCCGGATCCCTCGAGG  
GCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCT  
TCTAGTTGCCAGCCATCTGTTGCTGGCCCTCCCCGTGCCTCCTGACCTGGAAAGGT  
GCCACTCCACTGTCCTTCATAAAAATGAGGAAATTGCATCGCATTGCTGAGTAGG  
TGTCTATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC  
AATAGCAGGCATGCTGGGGATGGGTGGCTATGGCTTCTGAGGCCGAAAGAACCA  
TGCCCAGTCATGCCAACGCTCTCCACCCAGCGCGGAGAACCTGCGTGAATCC  
ACTGGGGCGCG

**Fig. 17**

**pCR-XL-TOPO-CMV-PUR-attB (SEQ ID NO: 10)**

AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCTGGCCGATTCAATTAAATGCAGCTGGC  
ACGACAGGTTTCCCGACTGGAAAGCGGCAGTGAGCGAACGCAATTAAATGTGAGTTAGC  
TCACTCATTAGGCACCCCAGGCTTACACTTTATGCTTCCGGCTGTATGTTGTGGAA  
TTGTGAGCGGATAACAATTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTAT  
TTAGGTGACGCGTTAGAATACTCAAGCTATGCAAGCTTGGTACCGAGCTGGATCCA  
CTAGTAACGGCCAGTGTGCTGGAAATTGCCCTGGCCGAATAAAATATCTTATTT  
TCATTACATCTGTGTTGGTTTTGTGAAATCGATAGTACTAAACATACGCTCTCCAT  
CAAAACAAAAGAAACAAAACAGCTTACGCGCTTAGCCCTCGAGCAGG  
GCCAGAACATTCTCATCGATAGGTACCGAGCTTACGCGTGTAGCCCTCGAGCAGG  
ATCTACATATTGAACTAAATATTGGCAATTAGCCATTAGTCATTGGTTATATAGCATAA  
ATCAATATTGGCTATTGGCCATTGCAACGTTGACATTGATTAGTACAGTTAATAG  
TAATCAATTACGGGGTCAATTAGTCATAGCCATATATGGAGTCCCGTACATAACTT  
ACGGTAAATGGCCCGCTGGCTGACCGCCAACGACCCCCGCCATTGACGTCAATAATG  
ACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAATGGGTGGAGTAT  
TTACGGTAAACTGCCACTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT  
ATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCAGTACATGACCTTACGG  
GACTTTCCCTACTTGGCAGTACATCTACGTTAGTCATCGCTTACGGTGTAGGGAGT  
TTTGGCAGTACATCAATGGGCGTGGATAGGGTTTGACTCACGGGATTTCAGTCTC  
CACCCCATTGACGTCATGGGAGTTTGGCACCAGGAACTAACGGGACTTCCAAAAA  
TGTGTAACAACACTCGCCCCATTGACGAAATGGCGGTAGGCGTGTACGGTGGGAGGTC  
TATATAAGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGT  
TTGACCTCCATAGAACGACACGGGACCGATCCAGCCTCCCCCTGAAGCTGACTCTAGG  
GGCTCGAGATCTGCGATCTAAGTAAGCTTGCATGCCCTGAGGTGGCCGCCACGACCGGT  
GCCGCCACCATCCCCCTGACCCACGCCCTGACCCCTCACAGGAGACGCCCTTCCATGAC  
CGAGTACAAGCCCACGGTGCCTGCCACCCGCGACGACGGTCCCCCGGCCGTACGCAC  
CCTGCCGCCGCGTCTGCCGACTACCCCGCACGCCACCGTCGACCCGGACCGCCA  
CATCGAGCGGGTCAACGGAGCTGAAAGAACACTTCTCACGCCGCTCGGGCTCGACATCGG  
CAAGGTGTTGGGTCGGGACGAGCGCCGGTGGCGCTGGACCACGCCGGAGAGCGT  
CGAAGCGGGGGCGGTGTTGCCGAGATCGGCCGCGCATGGCGAGTTGAGCGGTTCCCG  
GCTGCCGCGCAGAACAGATGGAAGGCCCTCTGGCGCCGACCGGCCAACGGGCAAGGG  
GTGGTTCTGCCACCGTCTGCCGACCGACCAGGGCAAGGGTCTGGCAGCGC  
CGTCGTGCTCCCCGGAGTGGAGGCCGAGCGGCCGGGGTGCCCTTCTGGAGAC  
CTCCGCCCGCCGCAACCTCCCCCTCTACGAGCGGCTCGGCTTACCGTCACCGCGACGT  
CGAGGTGCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGGCCGGTGCCTGACGCC  
GCCCAACGACCCGAGGCCGACCGAACAGGAGCGCACGCCATGGCTCCGACCGAAG  
CCGACCCGGGCCGCCGACCCGACCCGCCGGGGCACCGACTCTAGAGTC  
GGGGCGGGCGGCCGCTCGAGCAGACATGATAAGATAATTGATGAGTTGGACAAACCA  
CAACTAGAATGAGTGGAAAAAAATGCTTATTGTAAGGATGCTATTGCTTATGCTTAT  
TTGTAACCAATTAAAGCTGCAATAAACAAAGTAAACAAACAAATTGCAATTCAATTATG  
TTCAGGTTCAAGGGGAGGTGTTGGAGTTTTAAAGCAAGTAAACCTCTACAAATG  
GTAACATGATAAGGATCAATTGGCTCAGGTACCGTCGACGATGTAGGTACGGTCTC  
GAAGCGCGGTGCGGTGCCAGGGCGTGCCTGGCTCCCGGGCGCGTACTCCACCTC  
ACCCATCTGGTCCATCATGATGAAACGGGTGAGGTGGCGGTAGTTGATCCGGCGAACGC  
GCCGCCACCGGGAAAGCCCTGCCCTGAAACCGCTGGCGCGTGGTCACGGTAGC  
GGGACGTGCGACGGCGTCGGCGGTGCGGATACCGGGGCGAGCGTCAGCGGTTCTCGAC  
GGTCACGGCGGGCATGTCGACAGCGAATTGATCCGTCGACCGATGCCCTTGAGAGCGCTT  
CAACCCAGTCAGCTCTTCCGGTGGCGCGGGCATGACTATCGTCGCCGACCTTATGAC  
TGTCTCTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTTCCGCTCTCGC  
TCACTGACTCGCTCGCTCGGTGTTGGCTCGGGCGAGCGGTACAGCTCACTCAAAGG  
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGAGGAAAGAACATGAAGGGCGAAT  
TCTGCAGATATCCATCACACTGGCGCGCTCGAGCATGATCTAGAGGGCCAATTGCG  
CCTATAGTGAAGTCGTTACAATTCACTGGCGCGTGTGTTACAACGTCGTGACTGGAAA  
ACCCCTGGCGTTACCCAACCTTAATGCCCTGCAAGCACATCCCCCTTCCGCCAGCTGGCGTA  
ATAGCGAAGAGGCCGACCGATGCCCTCCAAACAGTTGCGCAGCCTACGTACGGC  
AGTTAAGGTTACACCTATAAAAGAGAGAGGCCGTTATCGTCTGTTGTGGATGTACAGA  
GTGATATTATTGACACGCCGGCGACGGATGGTACCGTCTGGCCAGTGCACGTCTGC

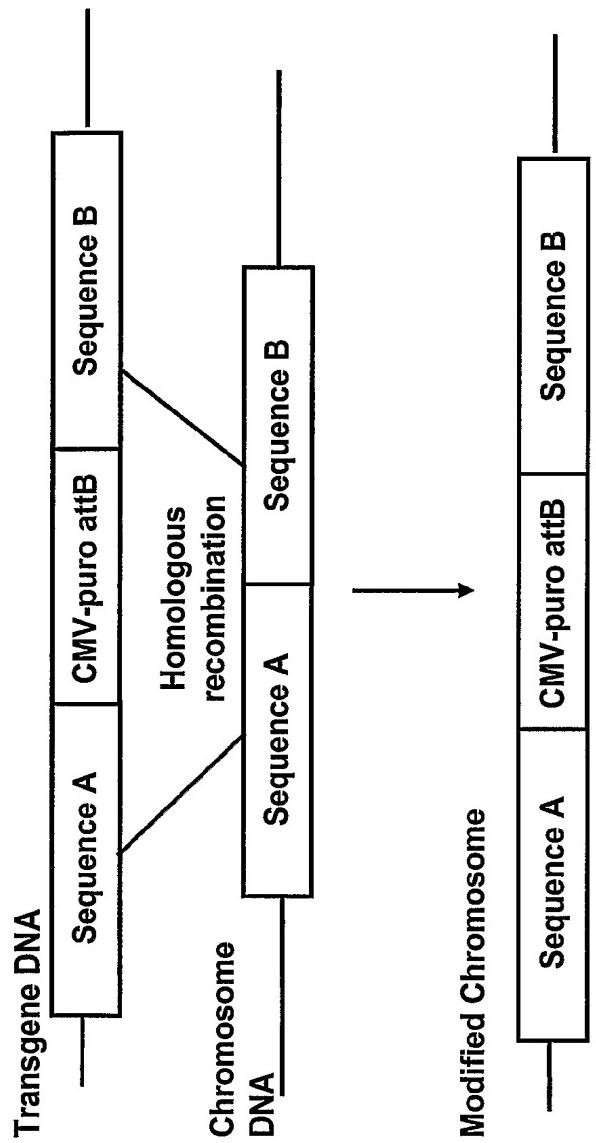
TGT CAGATAAAAGTCTCCGTGAACCTTACCGGGTGGTGCATATCGGGGATGAAAGCTGGC  
GCATGATGACCACCGATATGCCAGTGTGCCGGTCTCCGTTATCGGGGAAAGAAGTGGCTG  
ATCTCAGCCACCGCGAAAATGACATCAAAACGCCATTAACCTGATGTTCTGGGGAAATAT  
AAATGTCAGGCATGAGATTATCAGGGATCTCACCTAGATCCTTCACGTAGAAAG  
CCAGTCCGCAGAACCGGTGCTGACCCGGATGAATGTCAGCTACTGGGCTATCTGGACAA  
GGGAAAACGCAAGCGAAAGAGAAAGCAGGTAGCTTGCACTGGGCTTACATGGCGATAGC  
TAGACTGGCGGTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGCGCCCTG  
GTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTCTCGGCCCAAGGATCTGAT  
GGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTCGCATGATTGAAAC  
AAGATGGATTGCACCGCAGGTTCTCCGGCGCTTGGGTGGAGAGGCTATTGGCTATGACT  
GGGCACAACAGACAATCGGCTGCTGTGATGCCGCGTGTTCGGCTGTCAGCGCAGGGC  
GCCCGGTTCTTTGTCAAGACCGACCTGTCCGGTGCCTGAAATGAACTGCAAGACGAGG  
CAGCGCGGCTATCGGGCTGGCACGACGGCGTCTTGCAGCTGTGCTCGACGTTG  
TCACTGAAGCGGAAGGGACTGGCTGCTATTGGCGAAGTGCCGGGCAGGATCTCCTGT  
CATCTCACCTTGCTCCTGCCAGAAAGTATCCATATGGCTGATGCAATGCCGGCTGC  
ATACGCTTGATCCGGCTACCTGCCATTGACCAAGCGAAACATCGCATCGAGCGAG  
CACGTACTCGGATGGAAGCGGTCTGCGCAGGCTCAAGGCAGGATGCCGACGGCGAGGATC  
TCGTCGTGACCCATGGCGATGCCGCTTGCGAATATCATGGGAAATGGCCGCTTT  
CTGGGATTCATCGACTGTGGCCGCTGGGTGTGGGGACCGCTATCAGGACATAGCGTTGG  
CTACCCGTGATATTGCTGAAGAGCTTGGGGCGAATGGGCTGACCGCTTCTCGTGTGTT  
ACGGTATGCCGCTCCGATTGCAAGCCATGCCCTTATCGCCTTCTGACGAGTTCT  
TCTGAATTATTAAACGCTTACAATTCTGATGCGTATTTCCTTACGCACTGTGCG  
GTATTTCACACCGCATAACAGGTGGCACTTTGGGGAAATGTGCGCGAACCCCTATTG  
TTTATTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCCTGATAAAT  
GCTTCATAATAGCACGTGAGGAGGGCCACCATGCCAAGTTGACAGTGCCGTTCCGGT  
GCTCACCGCGCGACGTCGCCGGAGCGGTGAGTTCTGGACCGACGGCTCGGGTCTC  
CCGGGACTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGACGACGTGACCCCTGTTCAT  
CAGCGCGGTCCAGGACCAGGTGGTCCGGACAACACCCCTGCCCTGGGTGTGGTGC  
CCTGGACGAGCTGTACGCCAGTGGTCCGGAGGTCGTGTCAGCAACTCCGGGACGCC  
CGGGCCGGCATGACCGAGATGGCGAGCAGCGTGGGGGGAGTTGCCCTGCGCGA  
CCC GGCGCAACTCGTGCACCTCGTGGCCGAGGAGCAGGACTGACACGTGCTAAACT  
TCATTTTAATTAAAAGGATCTAGGTGAAGATCTTGTATAATCTCATGACCAAAAT  
CCCTTAACGTGAGTTCTGTTCCACTGAGCGTCAAGACCCCTAGAAAAGATCAAAGGATC  
TTCTTGAGATCCTTTCTGCGCTAATCTGCTGCTGCAAACAAAAACCCACCGCT  
ACCAGCGGTGGTTGCTGCCGATCAAGAGCTACCAACTCTTCCGAAGGTAACCTGG  
CTTCAGCAGAGCGCAGATAACCAAATACTGTCCTCTAGTGTAGCCGTAGTTAGGCCACCA  
CTTCAAGAAGCTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCTGTACCGAG  
TGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTGGACTCAAGACGGATAGTTACCGGA  
TAAGGCGCAGCGGTGGCTGAACGGGGGGTCTGTCACACAGCCCAGCTGGAGCGAAC  
GACCTACACCGAAGTACGAGATACTACAGCGTGAAGCTATGAGAAAGGCCACGCTCCGA  
AGGGAGAAAGCGGACAGGTATCCGTAAGCGGCAAGGGTGGAAACAGGAGAGCGCACGAG  
GGAGCTTCAAGGGGAAACGCCCTGGTATCTTATAGTCCTGCTGCCGCTTCC  
ACTTGAGCGTGCATTGGATGCTCGTCAGGGGGGGAGCCTATGGAAAACGCCAG  
CAACCGCGCCCTTTACGGTCTGGCTTTGCTGGCCTTTGCTCACATGTTCTTCC  
TGCCTATCCCCCTGATTCTGTTGATAACCGTATTACCGCCTTGTAGTGAGCTGATAACCGC  
TCGCCGAGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAG

**FIG. 18**

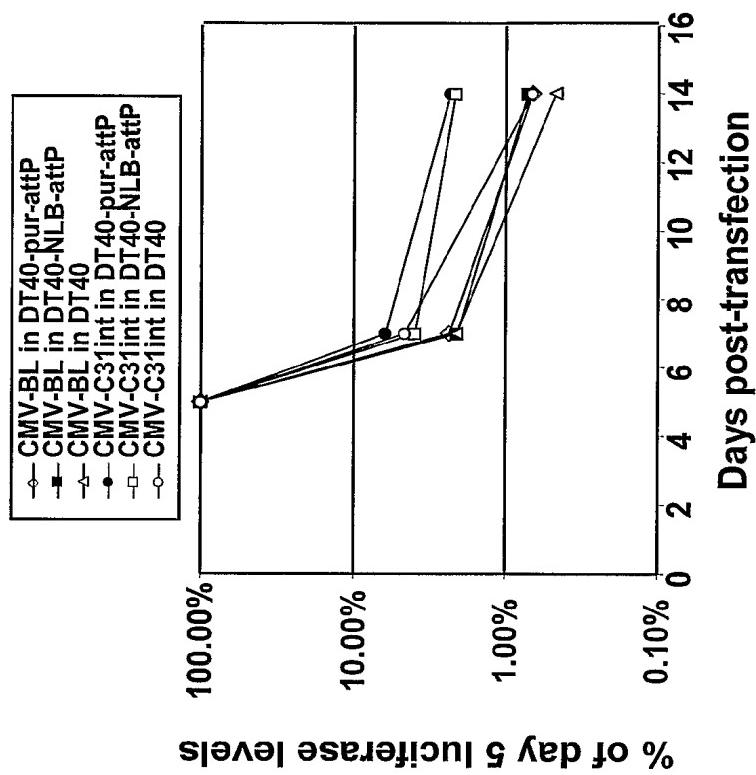
**SEQ ID NO: 11**

GACTAGTACTGACGGACACACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCGGGGCTT  
CACGTTTCCCAGGTCAAGCGGTTTCGGGAGTAGTGCCCCAACTGGGTAACCTTGAG  
TTCTCTCAGTTGGGGCGTAGGGTCGCCGACATGACACAAGGGTTGTGACCGGGTGGACA  
CGTACGCGGTGCTTACGACCGTCAGTCGCGGAGCGCGACTAGTACA

*Fig. 19*



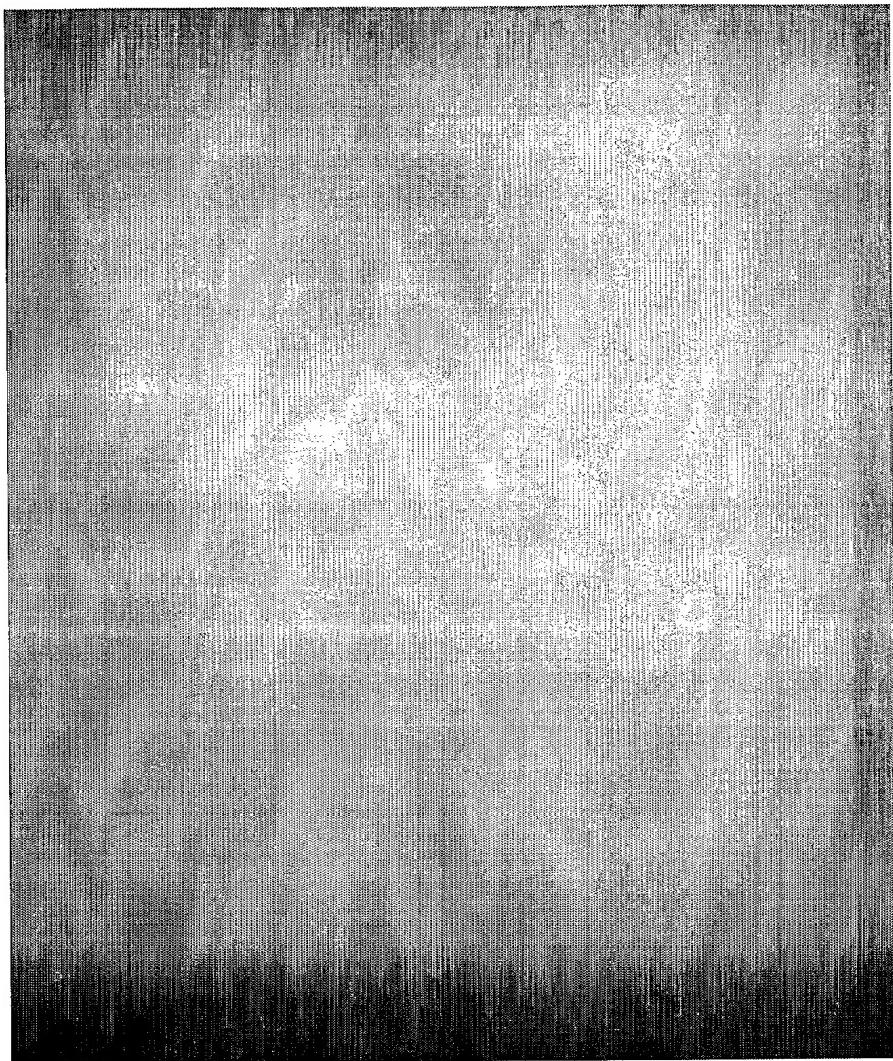
*Fig. 20*



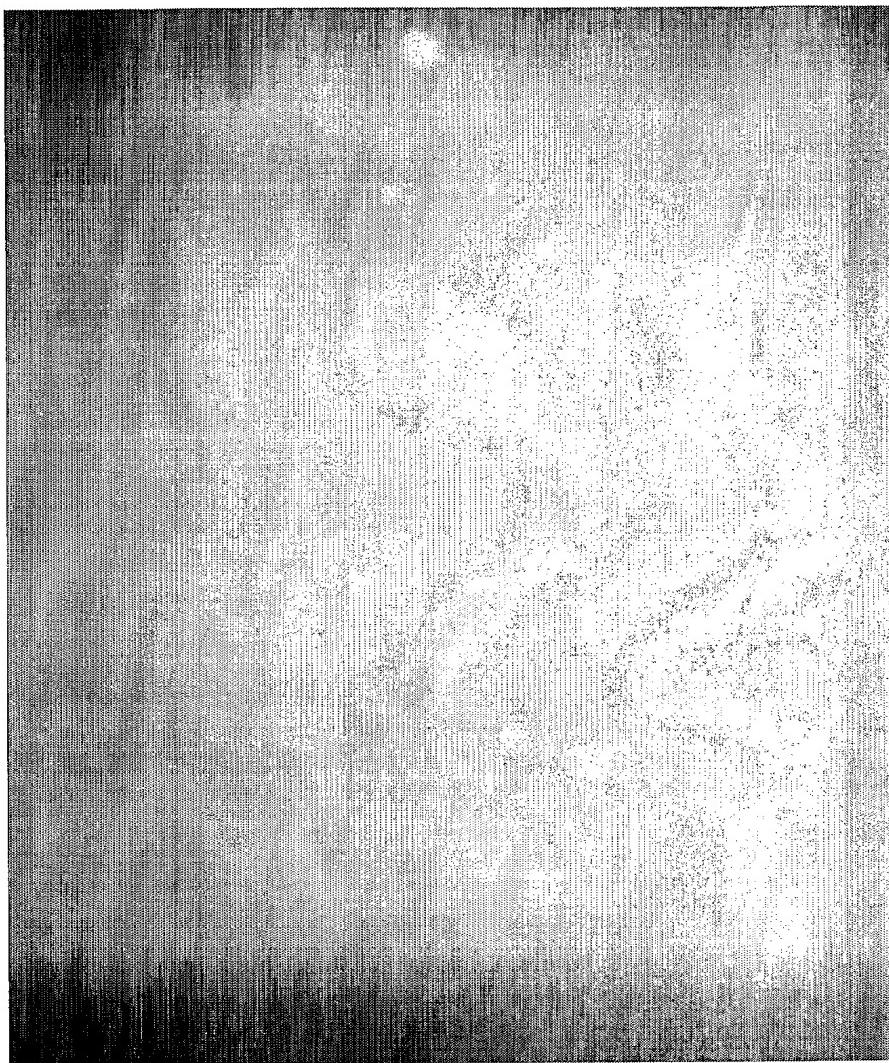
*Fig. 21*



*Fig. 22*



*Fig. 23*



*Fig. 24*

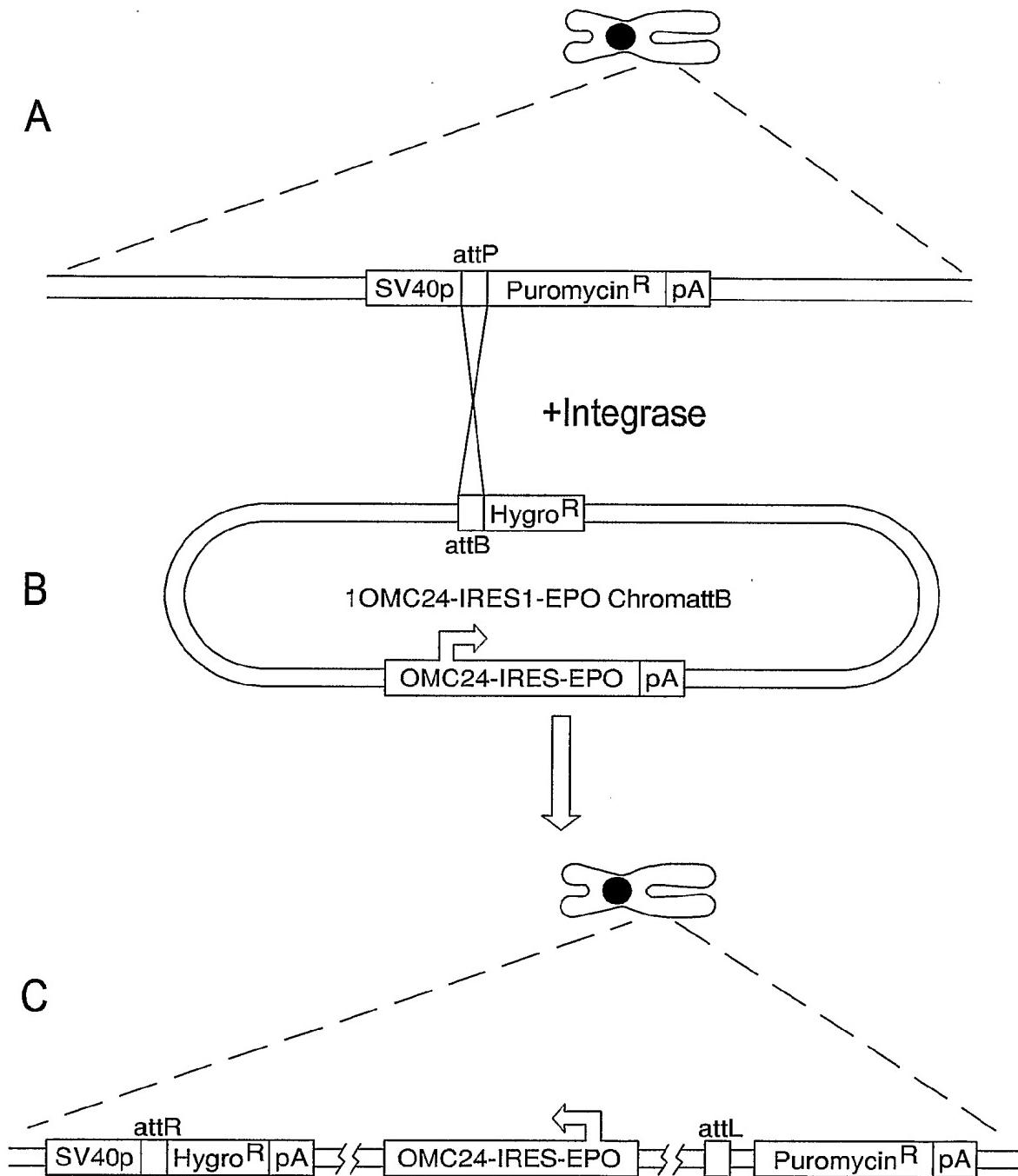


FIG. 25

025CIP4 PCT SEQ List.txt  
SEQUENCE LISTING

<110> AviGenics, Inc  
<120> MODIFICATION OF THE GENOME BY INTRODUCTION OF LARGE NUCLEIC ACID MOLECULES  
<130> AVI-025CIP4PCT  
<160> 13  
<170> PatentIn version 3.2  
<210> 1  
<211> 6230  
<212> DNA  
<213> Artificial Sequence  
<220>  
<223> Plasmid pCMV-31int  
<400> 1  
cattcccat tcaggctgcg caactgttgg gaagggcgat cggcgccggc ctcttcgcta 60  
ttacgccagc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatgcagg 120  
atcgatccag acatgataag atacattgtat gagtttgac aaaccacaac tagaatgcag 180  
tgaaaaaaaaat gctttatTTT tgaaatttgt gatgctatttgc ctttatttgt aaccattata 240  
agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg 300  
gaggtgtggg aggtttttta aagcaagtaa aacctctaca aatgtggat ggctgattat 360  
gatcatgaac agactgtgag gactgagggg cctgaaatga gccttggac tgtgaatcta 420  
aaatacacaa acaatttagaa tcactagctc ctgtgtataa tattttcata aatcatactc 480  
agtaagcaaa actctcaagc agcaagcata tgcagctagt ttaacacatt atacacttaa 540  
aaatTTTata tttaccttag agctttaaat ctctgttagt agtttgcatttca attatgtcac 600  
accacagaag taaggttcct tcacaaagat cccaaagctag cttataatac gactcactat 660  
agggagagag ctatgacgac gcatgcacgc gtaagcttgg gcccctcgag ggatccgggt 720  
gtctcgctac gccgctacgt cttccgtgcc gtcctggcg tcgtcttcgt cgctcgctgt 780  
cgccggcttc gcccacgtga tcgaagcgac cttctcgatg ggcgttccct gccccctgcc 840  
cgtagtcgac ttctgtacaa cgatcttgc tacgaagagc ccgacgaaca cgcgttgc 900  
gtctactgac gcgcgcgcgc accacgactt agggccggc gggtcagcgt cggcgtcttc 960  
ggggaaaccat tggtcaaggga aagcttcgg ggcttcggcg gctcaagtt cggcaagccg 1020  
ctcttcggcc ctttgctgcc ggagcgtcag cgctgcgttgc tgcttccggaa agtgcatttcc 1080  
gccaacgggt ccgtcgatc cgccgtccgc gcggtcttcg tacagcttttcaaggcggtt 1140  
cagggcgctcg gcgcgctccg caacaagggtt cggccgttgc cggctttct caggcgccctc 1200  
agtgagcttgc cgaagcgac gggcggttc ccacagaagc gccaacgtct cttcgctgcc 1260

## 025CIP4 PCT SEQ List.txt

ttcggcgtgc ctgatcttgt tgaagatgcg ttccgcaacg aacttgcga gtgccgccat	1320
gctgacgttg cacgtgcctt cgtgctgccc aggtgcggac gggtcgacca cttccggcg	1380
acggcagcgg taagagtccct tgatcgattc ttccccgcgc ttcgaagtca tgacggcgcc	1440
acaactcgca tacagcttgt ccatggcgga cagaatggct tgccccggg aaagccccctt	1500
gccgcgcccc ctgccgtcca accacgcctg aagctcatac cactcagcgg gctcgatgat	1560
cggtccgcaa tcaagctcgaa ccggccggag cgtgatcggg tcgcgctgaa tgcggttaacc	1620
ctcaatcttc gtggtcggcg tgccgtccgg cttcttctt tagatcacct cagcggcgaa	1680
gcccgcataa cgcggtccc gaaggattcg cataacggtt gccgggtccc aggcgcttga	1740
agcggtcttc ttcccaatcg tctcgccccg ggtcggcacg gcgtcagcgt ccatgcgcctt	1800
acaaaagcccc gtgatgctgc ccgggtgaat ggcggcttga ctgcccggct tgaagggaaag	1860
gtgtttgtgc gtcttgatct cacgcccacca ccaccggatt acgtcgggct cgaactcgaa	1920
gggtccggta aggggagtgg tcgagtgcgc aagcttgtt atgacgacat tgaccattcg	1980
gccgttgcgc gtgatctcct tcgtctccga aacaagctcg aagccgtaag gcgccttccc	2040
gccgacgtac ccgcccatt cgcgctgaag gttttcgatc tcgagaatct tcgcccactt	2100
cagcgaagat tctttgtgcg acgcgtcgag ccgcataatc aggtgaatca ggtccatgac	2160
gtttccctgc cggaagacgc cttcctgagt gaaaacaatc gtcacgcccc gggcgagcaa	2220
ttccgagaca atcggaatcg cgtccatgac cttcaggcgc gagaagcgcg acacgtcata	2280
gacaatgatc atgtttagcc gcccggcgcg gcattcggtt aggatgcgtt cgaactccgg	2340
gcgcctccgccc gtcccgaaacg ccgacgtgcc cggcgcttcg ctgaaatgcc cgacgaacct	2400
gaaccggccc ccgtcgcgct cgacttcgcg ctgaaggctcg gccgccttgt cttcggttggc	2460
gctacgctgt gtcgctggc ttgctgcgt cgaattctcg cgctcgcg actgacggtc	2520
gtaagcaccc gcgtacgtgt ccaccccggt cacaacccct tgtgtcatgt cggcgaccct	2580
acgactagtg agtcgtcga cccggaaatt cggaccggt acctgcaggc gtaccttcta	2640
tagtgtcacc taaatagctt tttgcaaaag cctaggctag agtccggagg ctggatcggt	2700
cccggtgtct tctatggagg tcaaaacagc gtggatggcg tctccaggcg atctgacgg	2760
tcactaaacg agctctgctt atatagaccc cccaccgtac acgcctaccg cccatttgcg	2820
tcaatggggc ggagttgtta cgacatttg gaaagtcccg ttgattttgg tgccaaaaca	2880
aactcccatt gacgtcaatg gggtgagac ttggaaatcc ccgtgagtc aaccgctatc	2940
cacgcccatt gatgtactgc caaaaccgca tcaccatggt aatagcgatg actaatacg	3000
agatgtactg ccaagtagga aagtccata aggtcatgta ctggcataa tgccaggcgg	3060
gccatttacc gtcattgacg tcaatagggg gcgtacttgg catatgatac acttgcgtt	3120

025CIP4 PCT SEQ List.txt

ctgccaagtg	ggcagttac	cgtaaatact	ccaccattg	acgtcaatgg	aaagtcccta	3180
ttggcggtac	tatgggaaca	tacgtcatta	ttgacgtcaa	tgggcggggg	tcgttggcg	3240
gtcagccagg	cgggccattt	accgtaaagtt	atgtaaacgac	ctgcacgatg	ctgtttcctg	3300
tgtgaaattg	ttatccgctc	acaattccac	acattatacg	agccggaagc	tataaagtgt	3360
aaagcctggg	gtgcctaatg	agtgaaaggg	cctcgtatac	gcctatTTT	ataggttaat	3420
gtcatgataa	taatggtttc	ttagacgtca	ggtggcactt	ttcggggaaa	tgtgcgcgga	3480
acccctattt	gtttatTTT	ctaaatacat	tcaaatatgt	atccgctcat	gagacaataa	3540
ccctgataaa	tgcttcaata	atattgaaaa	acgcgcgaat	tgcaagctct	gcattaatga	3600
atcggccaac	gcgcggggag	aggcggttg	cgtattggc	gctttccgc	ttcctcgctc	3660
actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3720
gtaatacgtt	tatccacaga	atcagggat	aacgcaggaa	agaacatgtg	agcaaaaggc	3780
cagcaaaagg	ccaggaacct	taaaaaggcc	gcgttgcgtgg	cgttttcca	taggctccgc	3840
ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcggaa	cccgacagga	3900
ctataaagat	accaggcgTT	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	3960
ctgcccctta	ccggataacct	gtccgcctt	ctcccttcgg	gaagcgtggc	gctttctcaa	4020
tgctcacgct	gtaggtatct	cagttcggtg	tagtcgttc	gctccaagct	gggctgtgtg	4080
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tctttagtcc	4140
aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	4200
gcgaggatag	taggcggtgc	tacagagttc	ttgaagtgg	ggcctaacta	cggctacact	4260
agaaggacag	tatTTggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	4320
ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggTTTTT	tgttgcaag	4380
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	4440
tctgacgctc	agtggAACGA	aaactcacgt	taagggattt	tggtcatgcc	ataacttcgt	4500
atagcataca	ttatacgaag	ttatggcatg	agattatcaa	aaaggatctt	cacctagatc	4560
cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	4620
gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgTTCA	4680
tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	4740
ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	4800
ataaaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	4860
atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcGCCAGT	taatagTTG	4920
cgcaacgttg	ttGCCATTGC	tacaggcatc	gtggtgtcac	gctcgtcgTT	tggtatggct	4980
tcattcagct	ccggTTCCCA	acgatcaagg	cgagttacat	gatccccat	gttgtgcAAA	5040

## 025CIP4 PCT SEQ List.txt

aaagcggtta	gctccttcgg	tcctccgatc	gttgtcagaa	gtaagttggc	cgcagtgtta	5100
tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	5160
ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	5220
agttgctctt	gcccgccgtc	aatacggat	aataccgcgc	cacatagcag	aactttaaaa	5280
gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	5340
agatccagtt	cgatgtAACCC	cactcgtgca	ccccaaactgat	cttcagcatc	ttttactttc	5400
accagcgTTT	ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaaa	gggaataagg	5460
gcgacacgga	aatgttgaat	actcataactc	ttcccttttc	aatattattt	aagcatttat	5520
cagggttatt	gtctcatgcc	aggggtgggc	acacatattt	gataccagcg	atccctacac	5580
agcacataat	tcaatgcgac	ttccctctat	cgcacatctt	agacctttat	tctccctcca	5640
gcacacatcg	aagctgccg	gcaagccgtt	ctcaccagtc	caagacctgg	catgagcgga	5700
tacatattt	aatgttattt	aaaaataaa	caaatagggg	ttccgcgcac	atttccccga	5760
aaagtgccac	ctgaaattgt	aaacgttaat	atttgtttaa	aattcgcgtt	aaatttttgt	5820
taaatcagct	catttttaa	ccaataggcc	gaaatcgca	aaatccctta	taaatcaaaa	5880
gaatagaccg	agatagggtt	gagtgttgg	ccagtttgg	acaagagtcc	actattaaag	5940
aacgtggact	ccaaacgtcaa	agggcgaaaa	accgtctatc	agggcgatgg	cccactacgt	6000
gaaccatcac	cctaattcaag	tttttgggg	tcgagggtgcc	gtaaagact	aaatcggaac	6060
cctaaaggga	gcccccgtt	tagagcttga	cggggaaagc	cggcgttacgt	ggcgagaaag	6120
gaagggaaga	aagcgaaagg	agcggcgct	agggcgctgg	caagtgttagc	ggtcacgctg	6180
cgcgttaacca	ccacacccgc	cgcgtttaat	gcgcccgtac	agggcgcg	tc	6230

<210> 2  
<211> 5982  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Plasmid pCMV-luc-attB

<400> 2	ctctatcgat	aggtaccgag	ctcttacgatc	tgcttagccct	cgagcaggat	ctatacattt	60
	aatcaatattt	ggcaatttgc	catatttagtc	attggttata	tagcataaat	caatattggc	120
	tattggccat	tgcatacgat	gtatctat	cataatatgt	acatttat	tggctcatgt	180
	ccaaatatgac	cgcctatgtt	acattgatta	ttgacttagtt	attaatagta	atcaattacg	240
	gggtcattag	ttcatagccc	atatatggag	ttccgcgtt	cataacttac	ggttaatggc	300
	ccgcctggct	gaccgccccaa	cgaccccccgc	ccattgacgt	caataatgac	gtatgttccc	360

025CIP4 PCT SEQ List.txt  
atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact 420  
gcccacttgg cagtacatca agtgttatcat atgccaagtc cgccccstat tgacgtcaat 480  
gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga ctttcctact 540  
tggcagtaca tctacgtatt agtcatcgct attaccatgg ttagtgcgggt ttggcagtac 600  
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac 660  
gtcaatggga gtttgttttgcaccaaaaat caacgggact ttccaaaatgcgtaaacaac 720  
tccgccccat tgacgcaaatttgcggtagg cgtgtacgggtt gggaggtcta tataaggcaga 780  
gctcgtttag tgaaccgtca gatcgccctgg agacgcccatttc cacgctgttt tgacctccat 840  
agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct 900  
gcgatctaag taagcttggc attccggtagc tgggttggaaa gccaccatgg aagacgcca 960  
aaacataaag aaaggccccgg cgccattctatccctggaa gatggaaaccg ctggagagca 1020  
actgcataag gctatgaaga gatacgccct ggttcctggaa acaattgctt ttacagatgc 1080  
acatatcgag gtggacatca cttacgctga gtacttcgaa atgtccgttc ggttggcaga 1140  
agctatgaaa cgatatgggc tgaataaaaa tcacagaatc gtcgtatgca gtgaaaactc 1200  
tcttcatttc tttatgccgg tgggtggcgc gttatattatc ggagttgcag ttgcgcggc 1260  
gaacgacatt tataatgaac gtgaatttgct caacagtatg ggcatttcgc agcctaccgt 1320  
gggttcgtt tccaaaaagg ggttgcaaaaa aattttgaac gtgaaaaaaaa agctcccaat 1380  
catccaaaaaa attattatca tggattctaa aacggattac cagggatttc agtcgtatgta 1440  
cacgttcgtc acatctcattc tacctccgg ttttaatgaa tacgattttg tgccagagtc 1500  
cttcgatagg gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc 1560  
taaagggtgtc gctctgcctc atagaactgc ctgcgtgaga ttctcgcatg ccagagatcc 1620  
tattttggc aatcaaatac ttccggatac tgcgattttt aagtgttgcattccatca 1680  
cggttttggaa atgtttacta cactcgata tttgatatgt ggatttcgag tcgtcttaat 1740  
gtatagattt gaagaagagc tggttctgag gagccttcag gattacaaga ttcaaagtgc 1800  
gctgctgggtg ccaaccctat tctccttctt cgccaaaagc actctgattt acaaatacga 1860  
tttatctaatttacacgaaa ttgcattctgg tggcgctccc ctctctaagg aagtgcgggaa 1920  
agcgggttgcc aagaggttcc atctgccagg tatcaggcaa ggatatgggc tcactgagac 1980  
tacatcagct attctgatta cacccgaggg ggatgataaa ccgggcgcgg tcggtaaagt 2040  
tggccattttttgaagcga aggttgcggatcc gggaaaacgc tgccgtttaa 2100  
tcaaagaggc gaactgtgtg tgagaggtcc tatgattatg tccgggtatg taaacaatcc 2160  
ggaagcggacc aacgccttga ttgacaagga tggatggcta cattctggag acatagctta 2220  
ctgggacgaa gacgaacact tcttcattcgatcc tggccgttgc aagtctctga ttaagtacaa 2280

## 025CIP4 PCT SEQ List.txt

aggctatcag	gtggctcccg	ctgaatttggaa	atccatcttgc	ctccaacacc	ccaacatctt	2340
cgacgcagg	gtcgcaagg	ttcccgacga	tgacgcccgt	gaacttcccg	ccgcccgttgt	2400
tgtttggag	cacggaaaga	cgatgacgga	aaaagagatc	gtggattacg	tcgcccagtca	2460
agtaacaacc	gcgaaaaagt	tgcgcggagg	agttgtgttt	gtggacgaag	taccgaaagg	2520
tcttaccgga	aaactcgacg	caagaaaaat	cagagagatc	ctcataaagg	ccaagaaggg	2580
cgaaaaagatc	gccgtgtaat	tctagagtcg	gggcggccgg	ccgcttcgag	cagacatgat	2640
aagatacatt	gatgagtttgc	gacaaaccac	aactagaatg	cagtaaaaaa	aatgctttat	2700
tttgtgaaatt	tgtgatgcta	ttgctttattt	tgttaaccatt	ataagctgca	ataaacaagt	2760
taacaacaac	aattgcatttc	attttatgtt	tcaggttcag	ggggaggtgt	gggaggtttt	2820
ttaaagcaag	taaaacctct	acaaatgtgg	taaaatcgat	aaggatcaat	tcggcttcag	2880
gtaccgtcga	cgatgttaggt	cacggctctcg	aagccgcgg	gcgggtgcca	ggcgtgccc	2940
ttgggctccc	cgggcgcgta	ctccacactca	cccatctgg	ccatcatgat	gaacgggtcg	3000
aggtggcggt	agttgatccc	ggcgaacgcg	cggcgcaccg	ggaagccctc	gccctcgaaa	3060
ccgctggcg	cggtggtcac	ggtgagcacg	ggacgtgcga	cggcgtcggc	gggtgcggat	3120
acgcggggca	gcgtcagcgg	gttctcgacg	gtcacggcgg	gcatgtcgac	agccgaatttgc	3180
atccgtcgac	cgtatcccc	gagagccccc	aacccagtca	gctcccccgg	gtgggcccgg	3240
ggcatgacta	tcgtcgccgc	acttatgact	gtcttcttta	tcatgcaact	cgtaggacag	3300
gtgccggcag	cgccttcccg	cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	3360
gcggcgagcg	gtatcagctc	actcaaaggc	ggtataacgg	ttatccacag	aatcagggga	3420
taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	3480
cgcgttgctg	gcgttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	3540
ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	3600
aagctccctc	gtgcgtctc	ctgttccgac	cctgccgtt	accggatacc	tgtccgcctt	3660
tctcccttcg	ggaagcgtgg	cgcttctca	atgctcacgc	tgttaggtatc	tcagttcggt	3720
gtaggtcggt	cgccttccaa	tggtgtgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	3780
cgccttatcc	ggtaactatc	gtcttgagtc	caacccggta	agacacgact	tatcgccact	3840
ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	gtaggcggtg	ctacagagtt	3900
cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	gtatggta	tctgcgtct	3960
gctgaagcca	gttaccttcg	aaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	4020
cgcgttgtagc	ggtggttttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	4080
tcaagaagat	ccttgatct	tttctacggg	gtctgacgct	cagtggAACG	aaaactcacg	4140

## 025CIP4 PCT SEQ List.txt

ttaagggatt ttggcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta	4200
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggctcg acagttacca	4260
atgcttaatc agtgaggcac ctatctcgc gatctgtcta tttcggtcat ccatagttgc	4320
ctgactcccc gtcgtgtaga taactacgt acgggagggc ttaccatctg gccccagtgc	4380
tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc	4440
agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat	4500
taattgttgc cggaagcta gagtaagtag ttgcgcaggta aatagttgc gcaacgttgt	4560
tgccattgct acaggcatcg tgggtgtcacg ctcgtcgtt ggtatggctt cattcagctc	4620
cggttcccaa cgatcaaggc gagttacatg atccccatg ttgtcaaaaa aagcggttag	4680
ctccttcggc cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatgg	4740
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac	4800
tggtgagtagt tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttgc	4860
cccgccgtca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat	4920
tggaaaacgt tcttcggggc gaaaactctc aaggatcttta ccgcgttga gatccagttc	4980
gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc	5040
tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa	5100
atgttgaata ctcatcttca atattattga agcattttatc agggttatttgc	5160
tctcatgagc ggatacatat ttgaatgtat tttagaaaaat aaacaaatag gggttcccg	5220
cacattttccc cgaaaagtgc cacctgacgc gcccgttagc ggccgtttaa gcgcggcgg	5280
tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gcccgtgc ccgcctttttc	5340
cgctttcttc ctttccttcc tcgcccacgtt cgccggctt ccccgtaag ctctaaatcg	5400
ggggctccct ttagggttcc gatttagtgc tttacggcac ctgcacccca aaaaacttga	5460
ttagggtgat ggttcacgtt gtggccatc gcccgtatag acggtttttc gccccttgac	5520
gttggagttcc acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc	5580
tatctcggtc tattcttttgc atttataagg gattttgccg atttcggcct attggtaaa	5640
aaatgagctg atttaacaaa aatttaacgc gaattttaaac aaaatattaa cgtttacaat	5700
ttcccatcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc	5760
gctattacgc cagcccaagc taccatgata agtaagtaat attaaggtac gggaggtact	5820
tggagcgccc gcaataaaaat atctttatcc tcattacatc tgtgtgttgg tttttgtgt	5880
gaatcgatag tactaacata cgctctccat caaaacaaaa cggaaacaaaa caaacttagca	5940
aaataggctg tccccagtc aagtgcaggt gccagaacat tt	5982

## 025CIP4 PCT SEQ List.txt

<210> 3  
<211> 5924  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Plasmid pCMV-luc-attP

<400> 3  
ctctatcgat aggtaccgag ctcttacgcg tgctagccct cgagcaggat ctatacattg 60  
aatcaatatt ggcaatttagc catattagtc attggttata tagcataaat caatattggc 120  
tattggccat tgcatacggt gtatctatat cataatatgt acatttatat tggctcatgt 180  
ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta atcaattacg 240  
gggtcattag ttcatalogccc atatatggag ttccgcgtta cataacttac ggtaaatggc 300  
ccgcctggct gaccgccccaa cgaccccccgc ccattgacgt caataatgac gtatgttccc 360  
atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact 420  
gcccacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat tgacgtcaat 480  
gacggtaaat ggcccgccctg gcattatgcc cagtacatga ctttacggga ctttcctact 540  
tggcagttaca tctacgtatt agtcatcgct attaccatgg ttagtgcgggtt ttggcagttac 600  
atcaatggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac 660  
gtcaatgggaa gtttgggttttgcacccaaaat caacgggact ttccaaaaatg tcgtaacaac 720  
tccgccttgcatt tgacgcaaattt gggcggtagg cgtgtacggt gggaggtcta tataaggcaga 780  
gctcggttag tgaaccgtca gatcgccctgg agacgccatc cacgctgttt tgacctccat 840  
agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct 900  
gcatctaaatg taagcttggc attccggtagc tggatggtaaa gccaccatgg aagacgcca 960  
aaacataaaatg aaaggccccgg cgccattctta tccgctggaa gatggaaaccg ctggagagca 1020  
actgcataatg gctatgaaga gatacgccct gtttcctggaa acaattgctt ttacagatgc 1080  
acatatcgag gtggacatca cttacgctga gtacttcgaa atgtccgttc gtttggcaga 1140  
agctatgaaa cgtatgggc tgaataaaaa tcacagaatc gtcgtatgca gtgaaaactc 1200  
tcttcatttgcattt tttatgccgg tggatggcgc gttatattatc ggagttgcag ttgcgcggc 1260  
gaacgacatt tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt 1320  
ggtgttcgtt tccaaaaagg gtttgcaaaa aatttgaac gtgaaaaaaa agctcccaat 1380  
catccaaaaaa attattatca tggattctaa aacggattac cagggatttc agtcgatgt 1440  
cacggttcgtc acatctcattc tacctccgg tttaatgaa tacgatgggt tgccagagtc 1500  
cttcgatagg gacaagacaa ttgcactgat catgaactcc tctggatcta ctggctgccc 1560  
taaagggtgtc gctctgcctc atagaactgc ctgcgtgaga ttctcgcatg ccagagatcc 1620

025CIP4 PCT SEQ List.txt

tatTTTggc	aatcaaatac	ttccggatac	tgcgattta	agtgttgttc	cattccatca	1680
cggTTTgga	atgtttacta	cactcgata	tttgatatgt	ggatttcgag	tcgtcttaat	1740
gtatagattt	gaagaagagc	tgTTTCTgag	gagcTTcag	gattacaaga	ttcaaagtgc	1800
gctgctggtg	ccaaccctat	tctcTTTcTT	cgccaaaagc	actctgattg	acaaatacga	1860
tttatctaata	ttacacgaaa	ttgcttctgg	tggcgctccc	ctctctaagg	aagtccccgg	1920
agcggTTGCC	aagaggTTCC	atctgccagg	tatcaggcaa	ggatatgggc	tcactgagac	1980
tacatcagct	attctgatta	cacccgaggg	ggatgataaa	ccgggCGCgg	tcggtaaagt	2040
tgttccattt	tttgaagcga	aggttgtgga	tctggatacc	ggggaaaacgc	tggcgTTaa	2100
tcaaagaggc	gaactgtgtg	tgagaggtcc	tatgattatg	tccggTTatg	taaacaatcc	2160
ggaagcGacc	aacgcTTga	ttgacaagga	tggatggcta	cattctggag	acatagctta	2220
ctgggacgaa	gacgaacact	tcttcatcgt	tgaccgcctg	aagtctctga	ttaagtacaa	2280
aggctatcag	gtggCTCCCg	ctgaattggta	atccatcttgc	ctccaacacc	ccaacatctt	2340
cgacgcaggt	gtcgcaggTC	ttcccgacga	tgacgcccgt	gaacttcccgt	ccgcccgttgt	2400
tgtttggag	cacggaaaga	cgtgacgga	aaaagagatc	gtggattacg	tcgcccagtca	2460
agtaacaacc	gcgaaaaaagt	tgcgcggagg	agttgtgttt	gtggacgaag	taccgaaagg	2520
tcttaccgga	aaactcgacg	caagaaaaat	cagagagatc	ctcataaaagg	ccaagaaggg	2580
cggaaagatc	gccgtgtaat	tctagagtgc	gggcggccgg	ccgcttcgag	cagacatgtat	2640
aagatacatt	gatgagTTTg	gacaaaccac	aactagaatg	cagtaaaaaa	aatgctttat	2700
ttgtgaaatt	tgtgatgcta	ttgctttatt	tgttaaccatt	ataagctgca	ataaacaagt	2760
taacaacaac	aattgcattc	atTTATGTT	tcaggttcag	ggggaggtgt	gggaggtttt	2820
ttaaaagcaag	taaaacCTCT	acaaatgtgg	taaaatcgat	aaggatcaat	tcggCTTCGA	2880
ctagtactga	cggcacacacc	gaagccccgg	cggcaaccct	cagcggatgc	cccggggctt	2940
cacgTTTCC	caggtcagaa	gcggTTTcg	ggagtagtgc	cccaactggg	gtAACCTTTG	3000
agttctctca	gttggggcg	tagggtcgcc	gacatgacac	aaggggTTgt	gaccgggggt	3060
gacacgtacg	cgggtgctta	cgaccgtcag	tcgcgcgagc	gcgactagta	caagccgaat	3120
tgatccgtcg	accgatgccc	ttgagagcct	tcaacccagt	cagctcTTc	cggTgggcgc	3180
ggggcatgac	tatcgTCGCC	gcacttatga	ctgtcttctt	tatcatgcaa	ctcgttaggac	3240
aggtGCCGGC	agcgCTCTTC	cgCTTCTcg	ctcactgact	cgctgcgctc	ggTCgtTCgg	3300
ctgcggcGAG	cggtatcagc	tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	3360
gataacgcag	gaaagaacat	gtgagaaaaa	ggccagcaa	aggccaggaa	ccgtaaaaag	3420
gcccgcgttgc	tggcTTTTT	ccataggctc	cgccccCTG	acgagcatca	caaaaatcga	3480
cgctcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	3540

## 025CIP4 PCT SEQ List.txt

ggaagctccc tcgtgcgtc tcctgttccg accctgccgc ttaccggata cctgtccgcc	3600
tttctccctt cgaaaagcgt ggcgtttct caatgctcac gctgttaggtta tctcagttcg	3660
gtgttaggtcg ttgcgtccaa gctggctgt gtgcacgaac cccccgttca gcccgaccgc	3720
tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca	3780
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag	3840
ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgt	3900
ctgctgaagc cagttacctt cgaaaaaaga gttggtagct cttgatccgg caaacaaacc	3960
accgctggta gcgggtggttt ttttgttgc aagcagcaga ttacgcgcag aaaaaaaagga	4020
tctcaagaag atcctttgat ctttctacg gggctgtacg ctcagtggaa cgaaaactca	4080
cgttaaggaa ttttggcat gagattatca aaaaggatct tcacctagat cttttaaat	4140
aaaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac	4200
caatgcttaa tcagtgaggc acctatctca gcgtatctgc tatttcgttc atccatagtt	4260
gcctgactcc ccgtcggtta gataactacg atacgggagg gcttaccatc tggccccagt	4320
gctgcaatga taccgcgaga cccacgctca cccggctccag attttatcagc aataaaccag	4380
ccagccggaa gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct	4440
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagtt gcgcaacgtt	4500
gttgccatttgc ctacaggcat cgtgggtgtca cgctcgctgt ttggatggc ttcattcagc	4560
tccggttccc aacgatcaag gcgagttaca tgatccccca tggtgtgcaa aaaagcggtt	4620
agctccttcg gtcctccgat cggtgtcaga agtaagttgg ccgcagtgtt atcactcatg	4680
gttatggcag cactgcataa ttctcttact gtcatgcccattt ccgtaaatgt ctttctgtg	4740
actggtagt actcaaccaa gtcattctga gaatagtgtt tgcggcgacc gagttgtct	4800
tgcccgccgt caatacggga taataccgcg ccacatagca gaactttaaa agtgcgtcatc	4860
attggaaaac gttttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt	4920
tcgatgtaac ccactcggtc acccaactga tcttcagcat cttttacttt caccagcggtt	4980
tctgggttagt caaaaacagg aaggcaaaat gccgaaaaa aggaaataag ggccgacacgg	5040
aaatgttcaa tactcataact ttccctttt caatattattt gaagcatttta tcagggttat	5100
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg	5160
cgcacatttc cccgaaaaagt gccacctgac gcgcctgtta gcggcgatt aagcgccggcg	5220
ggtgtggtagt ttacgcgcag cgtgaccgctt acacttgcca gcgccttagc gcccgtcct	5280
ttcgctttct tcccttcctt tctcgccacg ttgcggctt tccccgtca agctctaaat	5340
cgggggctcc ctttagggtt ccgatttagt gcttacggc acctcgaccc caaaaactt	5400

## 025CIP4 PCT SEQ List.txt

gattagggtg atggttcacg tagtggccca tcgcctgat agacggttt tcgccttgc	5460
acgttggagt ccacgttctt taatagtggc ctctgttcc aaactggAAC aacactcaac	5520
cctatctcggt tctattcttt tgatttataa gggatttgc cgatttcggc ctattggta	5580
aaaaatgagc tgatttaaca aaaatttaac gcgaattttt acaaaatatt aacgtttaca	5640
atttcccatt cgccatttcag gctgcgcaac tggggaaag ggcatcggt gcgggcctct	5700
tcgctattac gccagccaa gctaccatga taagtaagta atattaaggt acgggaggt	5760
cttggagcgg ccgcaataaaa atatctttat tttcattaca tctgtgtgtt gggttttgt	5820
gtgaatcgat agtactaaca tacgctctcc atcaaaacaa aacgaaacaa aacaaactag	5880
caaaataggc tgtccccagt gcaagtgcag gtgccagaac attt	5924

<210> 4  
<211> 5101  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Plasmid pCMV-pur-attB

<400> 4 ctagagtccgg ggcggccggc cgcttcgagc agacatgata agatacattt atgagtttg	60
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgtatgttat	120
tgctttatTT gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca	180
ttttatgttt caggttcagg gggaggtgtg ggaggTTTT taaagcaagt aaaacctcta	240
caaattgtgtt aaaatcgata aggatcaatt cggcttcagg taccgtcgac gatgttaggtc	300
acggtctcga agccgcggtg cgggtgccag ggcgtgcct tgggctcccc gggcgcgtac	360
tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggta gttgatcccg	420
gcgaacgcgc ggcgcaccgg gaagccctcg ccctcgaaac cgctggcgc ggtggtcacg	480
gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg	540
ttctcgacgg tcacggcggg catgtcgaca gccgaatttga tccgtcgacc gatgccttgc	600
agagccttca acccagtca gtccttccgg tgggcgcggg gcatgactat cgtcgccgca	660
cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc gctttccgc	720
ttcctcgctc actgactcgc tgcgctcggt cggtcggctg cggcggcgg tatcagctca	780
ctcaaaggcg gtaatacggt tatccacaga atcagggat aacgcaggaa agaacatgtg	840
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca	900
taggctccgc cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa	960
cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc	1020
tgttccgacc ctgcccgtta ccggatacct gtccgcctt ctccttcgg gaagcgtggc	1080

## 025CIP4 PCT SEQ List.txt

gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcggtc gctccaagct	1140
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg	1200
tcttgagtcc aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag	1260
gattagcaga gcgaggtatg taggcgggtgc tacagagttc ttgaagtgg tgcctaacta	1320
cggtacact agaaggacag tatttggat ctgcgctctg ctgaagccag ttaccttcgg	1380
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggttttt	1440
tgttgcaag cagcagatta cgccgagaaa aaaaggatct caagaagatc ctttgatctt	1500
ttctacgggg tctgacgctc agtggAACGA aaactcacgt taagggattt tggcatgag	1560
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat	1620
ctaaagtata tatgagtaaa cttggctga cagttaccaa tgcttaatca gtgaggcacc	1680
tatctcagcg atctgtctat ttctgtatc catagttgcc tgactccccg tcgtgttagat	1740
aactacgata cgggagggct taccatctgg ccccaagtgc gcaatgatac cgccgagaccc	1800
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag	1860
aagtggccct gcaactttat ccgcctccat ccagtctatt aatttgttgc gggaaagctag	1920
agtaagtagt tcgcccagtta atagttgcg caacgttgtt gccattgcta caggcatcgt	1980
ggtgtcacgc tcgtcggtt gtagggcttc attcagctcc ggttcccaac gatcaaggcg	2040
agttacatga tccccatgt tgtgaaaaa agcggttagc tccttcggc ctccgatcgt	2100
tgtcagaagt aagttggccg cagtgttatac actcatggtt atggcagcac tgcataattc	2160
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtagt caaccaagtc	2220
attctgagaa tagtgtatgc ggccgaccgag ttgctttgc ccggcgtcaa tacggataa	2280
taccgcgcca catagcagaa cttaaaagt gctcatcatt ggaaaacgtt cttcggggcg	2340
aaaactctca aggatcttac cgctgttgag atccagttcg atgtAACCCA ctcgtgcacc	2400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtagccaa aaacaggaag	2460
gcaaaatgcc gcaaaaaagg gaataaggc gacacggaaa tggtaatac tcataacttt	2520
ccttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt	2580
tgaatgtatt tagaaaaata aacaaatagg gttccgcgc acatttcccc gaaaagtgcc	2640
acctgacgctg ccctgttagcg ggcattaaag cgccgggggt gtggtggtt cgcgcagcgt	2700
gaccgctaca cttgccagcg ccctagcgcc cgctccttcc gctttcttcc ctcccttct	2760
cgccacgttc gcccgtttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg	2820
attttagtgct ttacggcacc tcgacccaa aaaacttgat tagggtgatg gttcacgtag	2880
ttggccatcg ccctgataga cggttttcg cccttgacg ttggagtcca cgttctttaa	2940

025CIP4 PCT SEQ List.txt

tagtggactc ttgttccaaa ctggAACACAC actcaACCCt atctcggtct attctttga	3000
tttataaggg attttgcgcga tttcggccta ttggtaaaaa aatgagctga tttAACAAAA	3060
atttAACGCG aattttAACA aaatattaac gtttacaatt tcccattcgc cattcaggct	3120
gCGCAACTGT tgGGAAGGGC gatcggtgcg ggcctttcg ctattacGCC agcccaagct	3180
accatgataa gtaagtaata ttaaggtacg ggaggtactt ggagcggccg caataaaata	3240
tctttatTTT cattacatct gtgtgttggt tttttgtgtg aatcgatagt actaacatac	3300
gctctccatc aaaacAAAAC gaaacAAAAC aaactagcaa aataggctgt ccccagtgcA	3360
agtgcaggtg ccagaacatt tctctatcga taggtaccga gctttacgc gtgctagccc	3420
tcgagcagga tctatacatt gaatcaatat tggcaattag ccatattagt cattggTTAT	3480
atagcataAAAt tcaaatATTgg ctattggcca ttgcatacgt tgtatctata tcataatATG	3540
tacatttATA ttggctcatg tccaatatga ccGCCatgtt gacattgatt attgactagt	3600
tattaatAGT aatcaattac ggggtcatta gttcatagcc catatatggA gttccgcgtt	3660
acataactta cggtAAATgg cccgcctggc tgaccGCCA acgACCCCCG cccattgacg	3720
tcaataatga cgtatgttcc catagtaacg ccaatAGGGA ctTTCCATTG acgtcaatgg	3780
gtggagttt tacggtaaac tgcccacttg gcagtagatc aagtgtatca tatGCCAAGT	3840
ccgcccccta ttgacgtcaa tgacgtaaaa tggccgcct ggcattatgc ccagtacatg	3900
accttacggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg	3960
gtgatgcgtt tttggcagta catcaatggg cgtggatAGC ggTTGACTC acggggattt	4020
ccaagtctcc accccattGA cgtcaatggg agtttGTTT ggcaccaAAA tcaacgggac	4080
tttccAAAt gtcgtAACAA ctccGCCCA ttgacgcaAA tggcggtag gcgtgtacgg	4140
tgggaggTCT atataAGCAG agctcgTTA gtgaACCgtc agatcgCCTG gagacGCCat	4200
ccacgctgtt ttgacctCCA tagaAGACAC cgggaccgat ccagcctCCC ctcgaagctc	4260
gactctaggG gctcgagatc tgcgatctAA gtaagcttgc atgcctgcag gtcggccGCC	4320
acgaccggTG ccGCCACCAT cccCTGACCC acgCCCTGA cccCTCACAA ggagacgacc	4380
ttccatgacc gagtacaAGC ccacggTgcG cctcgCCAcc cgcgacgacg tccccGGGc	4440
cgtacgcacc ctcGCCGCCG cgTTcgCCGA ctACCCGCC acgCGCCACA ccgtcgacCC	4500
ggaccGCCAC atcgagcGGG tcaccGAGt gcaAGAACtC ttccTcacGC gcgtcGGGct	4560
cgacatCGGC aaggTgtGGG tcgcggacGA cggGCCGCCG gtggcggTct ggaccacGCC	4620
ggagAGCGTC gaagcGGGGG cggTgttCGC cgagatCGGC ccgcgcAtgg ccgagttgag	4680
cggTTCCCGG ctggCCGCCG agcaacAGAt ggaAGGCCtC ctggCGCCGC accGGCCAA	4740
ggagCCGCCG tggttCCtgg ccaccgtcgg cgtctGCCc gaccaccagg gcaagggtct	4800
gggcagcGCC gtcgtgtcc ccggagtgga ggcggccgag cgcGCCGGG tgcccGCCtt	4860

## 025CIP4 PCT SEQ List.txt

cctggagacc tccgcgcccc gcaacacctcc cttctacgag cggtcggtc tcaccgtcac	4920
cgccgacgtc gaggtgcccc aaggaccgag cacctgggtgc atgaccgcga agcccgggtgc	4980
ctgacgccccg cccccacgacc cgtagcgccc gaccgaaagg agcgcacgac cccatggctc	5040
cgaccgaagc cgaccgggc ggcccccggc accccgcacc cgccccggag gcccacccgac	5100
t	5101

<210> 5  
<211> 5043  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Plasmid pCMV-pur-attP

<400> 5	
ctagagtccgg ggcggccggc cgcttcgagc agacatgata agatacattt atgagtttgg	60
acaaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgttat	120
tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca	180
ttttatgttt caggttcagg gggaggtgtg ggaggtttt taaagcaagt aaaacctcta	240
caaatgttgtt aaaatcgata aggatcaatt cggttcgac tagtactgac ggacacaccg	300
aagccccggc ggcaaccctc agcggatgcc ccggggcttc acgtttccc aggtcagaag	360
cggtttcgg gagtagtgcc ccaactgggg taaccttga gttctctcag ttggggcgt	420
agggtcgccg acatgacaca aggggttgtg accgggggtgg acacgtacgc gggtgcttac	480
gaccgtcagt cgcgcgagcg cgactagtac aagccgaatt gatccgtcga ccgtgcct	540
tgagagcctt caacccagtc agtccttcc ggtggcgcg gggcatgact atcgtcgccg	600
cacttatgac tgtcttcttt atcatgcaac tcgttaggaca ggtgccggca gcgtcttcc	660
gcttcctcgc tcactgactc gctgcgtcg gtcgttcggc tgcggcgagc ggtatcagct	720
cactcaaagg cggtataacg gttatccaca gaatcagggg ataacgcagg aaagaacatg	780
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcttttc	840
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga	900
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgtct	960
cctgttccga ccctgcccgt taccggatac ctgtccgcct ttctcccttc gggaaagcgtg	1020
gcgcgttctc aatgctcacg ctgttaggtat ctcagttcgg tgttaggtcgt tcgcgtccaag	1080
ctgggctgtg tgcacgaacc ccccggtcag cccgaccgtc gcgccttatac cgtaactat	1140
cgtctttagt ccaacccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac	1200
aggattagca gagcgaggtt tgttaggcgtt gctacagagt tcttgaagtg gtggcctaac	1260

025CIP4 PCT SEQ List.txt

tacggctaca	ctagaaggac	agtatTTggT	atctgcgctc	tgctgaagcc	agttaccttc	1320
ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtggttt	1380
tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaaggat	ctcaagaaga	tcctttgatc	1440
ttttctacgg	ggtctgacgc	tcagtggAAC	gaaaactcac	gttaagggat	tttggTCatG	1500
agattatcaa	aaaggatctt	cacctagatc	ctttaaatt	aaaaatgaag	ttttaaatca	1560
atctaaagta	tatATgagta	aactTggTCT	gacAGTTacc	aatGCTTAAT	cagtgaggca	1620
cctatctcag	cgatCTGTCT	atTCGTTCA	tccatAGTTG	cctgactccc	cgtcgtgtAG	1680
ataactacga	tacGGGAGGG	cttaccatct	ggccccAGTG	ctgcaatgat	accgcgagac	1740
ccacgctcac	cggCTCCAGA	tttatCAGCA	ataaaccAGC	cagCCGGAAG	ggccgagcgc	1800
agaagtggTC	ctgcaacttt	atCCGCCTCC	atCCAGTCTA	ttaattgttg	ccggGAAGCT	1860
agagtaagta	gttCGCCAGT	taatAGTTG	cgcaACGTTG	ttGCCATTGC	tacAGGCATC	1920
gtggTgtcac	gctcgtcgTT	tggTATGGCT	tcattcAGCT	ccggTTCCCA	acgatcaagg	1980
cgagttacat	gatCCCCCAT	gttGTGCAAA	aaAGCGGTTA	gctcCTTCGG	tcctCCGATC	2040
gttGTCAGAA	gtaAGTTGGC	cgcAGTGTta	tcactcatGG	ttatGGCAGC	actgcataat	2100
tctcttactg	tcatGCCATC	cgtaAGATGC	tttCTGTGA	ctggTgAGTA	ctcaaccaAG	2160
tcattCTGAG	aatAGTGTAT	gcggcGACCG	agttGCTCTT	gccCGGCGTC	aatacGGGAT	2220
aataccgcgc	cacatAGCAG	aactttAAA	gtgCTCATCA	ttggAAAACG	ttcttCGGGG	2280
cggAAAActCT	caaggatCTT	accGCTGTTG	agatCCAGTT	cgatGTAACC	cactCGTGC	2340
cccaactgat	cttcAGCATC	tttTACTTTC	accAGCGTT	ctgggtgAGC	aaaaacAGGA	2400
aggcaAAATG	ccgcaAAAAAA	gggaATAAGG	gcgacACGGA	aatGTTGAAT	actcataACTC	2460
ttcCTTTTC	aatattATTG	aagcATTAT	caggGTTATT	gtctcatGAG	cggatacata	2520
tttGAATGTA	tttagAAAAAA	taaACAAATA	ggggTTCCGC	gcacATTCC	ccgAAAAGTG	2580
ccacCTGACG	cggCCCTGTag	cggcgcatta	agcgcggcgg	gtgtggTGGT	tacgcgcAGC	2640
gtgaccGCTA	cactTGCCAG	cggCCCTAGCG	cccGCTCCTT	tcgCTTTCTT	ccCTTCCCTT	2700
ctcgCCACGT	tcgCCGGCTT	tccccgtCAA	gctctAAATC	ggggGCTCCC	tttagGGTTc	2760
cgatTTAGTG	cttTACGGCA	cctcGACCCC	aaaaAAACTG	attAGGGTGA	tggTcacGT	2820
agtggGCCAT	cggCCCTGATA	gacGGTTTT	cggCCtttGA	cgttggAGTC	cacGTTCTTT	2880
aatAGTGGAC	tctTGTtCCA	aactGGAACA	acactCAACC	ctatCTCGGT	ctattCTTT	2940
gatttATAAG	ggatTTGCC	gattTCGGCC	tattGGTTAA	aaaATGAGCT	gatttaACAA	3000
aaatttaACG	cgaattttAA	caaaatATTA	acgtttACAA	tttCCCATTC	gccattcagg	3060
ctgcgcAact	gttGGGAAGG	gcgatCGGTG	cggGCCTCTT	cgctattACG	ccagCCCAAG	3120
ctaccatGAT	aagtaAGTAA	tattaAGGTA	cggGAGGTAC	ttggAGCGGC	cgcaataAAA	3180

## 025CIP4 PCT SEQ List.txt

tatctttatt ttcattacat ctgtgttgtt gtttttgtt tgaatcgata gtactaacat	3240
acgctctcca tcaaaaacaaa acgaaacaaa acaaactagc aaaataggct gtccccagtg	3300
caagtgcagg tgccagaaca tttctctatc gataggtaacc gagctttac gcgtgcttagc	3360
cctcgagcag gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt	3420
atatacgata aatcaaatatt ggctattggc cattgcatac gttgtatcta tatcataata	3480
tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta	3540
gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg gagttccgcg	3600
ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc cgccattga	3660
cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat	3720
gggtggagta ttacggtaa actgcccact tggcagtaca tcaagtgtat catatgcca	3780
gtccgc(cc) tattgacgtc aatgacggta aatggccgc ctggcattat gccagtgaca	3840
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcatc gctattacca	3900
tggtgatgctg gtttggcag tacatcaatg ggcgtggata gcggtttgc tcacggggat	3960
ttccaagtct ccacccatt gacgtcaatg ggagtttggtt ttggcaccaa aatcaacggg	4020
actttccaaa atgtcgtaac aactccgccc cattgacgca aatggccgtt aggctgtac	4080
ggtggggaggt ctatataagc agagctcggt tagtgaaccg tcagatcgcc tggagacgcc	4140
atccacgctg tttgacctc catagaagac accgggaccg atccagcctc ccctcgaagc	4200
tcgactctag gggctcgaga tctgcgatct aagtaagctt gcatgcctgc aggtcggccg	4260
ccacgaccgg tgccgccacc atccccgtac ccacgccccct gacccctcac aaggagacga	4320
ccttccatga ccgagtacaa gcccacggtg cgccctgcca cccgcgacga cgtccccgg	4380
gccgtacgca ccctcgccgc cgcggtcgcc gactaccccg ccacgcgcca caccgtcgac	4440
ccggaccgccc acatcgagcg ggtcaccgag ctgcaagaac tcttcctcac gcgcgtcgaaa	4500
ctcgacatcg gcaagggtgtt ggtcgccggac gacggcccg cggtggccgt ctggaccacg	4560
ccggagagcg tcgaagcggg ggcgggtttc gccgagatcg gcccgccat ggccgagttt	4620
agcggttccc ggctggccgc gcagcaacag atgaaaggcc tcctggccgc gcaccggccc	4680
aaggagcccg cgtggttccct ggccaccggtc ggcgtctcgcc ccgaccacca gggcaagggt	4740
ctgggcagcg ccgtcggtct cccggagtg gaggccggccg agcgccggcgg ggtgcccgc	4800
ttcctggaga cctccgcgc cgcacaccctc cccttctacg agcggtcgcc tttcaccgtc	4860
accggccacg tcgaggtgcc cgaaggaccg cgacccgtt gcatgaccgg caagccgggt	4920
gcctgacgccc cgccccacga cccgcagcgc ccgaccgaaa ggagcgacg accccatggc	4980
tccgaccgaa gccgaccggc gggccccgc cgaccggca cccgcccccg aggcccaccg	5040

## 025CIP4 PCT SEQ List.txt

act

5043

<210> 6  
<211> 5041  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Plasmid pCMV-EGFP-attB  
  
<400> 6  
ctagagtcgg ggccgcggc cgcttcgagc agacatgata agatacattt atgagttgg 60  
acaaaccaca actagaatgc agtaaaaaaaa atgttttatt tgtaaaattt gtgtatgttat 120  
tgctttattt gtaaccatta taagctgcaa taaacaaggta aacaacaaca attgcattca 180  
ttttatgttt caggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta 240  
caaatgttgtt aaaatcgata aggtcaatt cggcttcagg taccgtcgac gatgttaggtc 300  
acggtctcga agccgcggtg cgggtgccag ggcgtgcctt tgggctcccc gggcgcgtac 360  
tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggta gttgatccc 420  
gcgaacgcgc ggcgcaccgg gaagccctcg ccctcgaaac cgctggcgc ggtggtcacg 480  
gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg 540  
ttctcgacgg tcacggcggg catgtcgaca gccgaattta tccgtcgacc gatgccctt 600  
agagccttca acccagtcag ctccctccgg tgggcgcggg gcatgactat cgtcgcggca 660  
cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc gctttccgc 720  
ttccctcgctc actgactcgc tgcgtcggt cggtcggtc cggcggcgg tatcagctca 780  
ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 840  
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttcca 900  
taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 960  
cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg tgcgtctcc 1020  
tgttccgacc ctgcccgtta ccggataacct gtccgcctt ctccctcgga gaagcgtggc 1080  
gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcggtc gctccaagct 1140  
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 1200  
tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 1260  
gattagcaga gcgaggtatg taggcgggtgc tacagagttc ttgaagtggg ggcctaacta 1320  
cggtcacact agaaggacag tattttgtat ctgcgtctg ctgaagccag ttaccttcgg 1380  
aaaaagagtt ggtagcttt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 1440  
tgtttgcaag cagcagatta cgccgcagaaa aaaaggatct caagaagatc ctttgatctt 1500  
ttctacgggg tctgacgctc agtggAACGA aaactcacgt taaggattt tggcatgag 1560

025CIP4 PCT SEQ List.txt

attatcaaaa aggatcttca cctagatcct tttaaatcaa aaatgaagtt tttaaatcaat	1620
ctaaagtata tatgagtaaa ctgggtctga cagttaccaa tgcttaatca gtgaggcacc	1680
tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgttagat	1740
aactacgata cgggagggct taccatctgg ccccagtgc gcaatgatac cgcgagaccc	1800
acgctcacccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag	1860
aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaaagctag	1920
agtaagtagt tcgccagttt atagttgcg caacgttgc gccattgcta caggcatcgt	1980
ggtgtcacgc tcgtcgttt gtagggcttc attcagctcc gggtcccaac gatcaaggcg	2040
agttacatga tccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt	2100
tgtcagaagt aagttggccg cagtgttac actcatggtt atggcagcac tgcataattc	2160
tcttactgtc atgccccatccg taagatgctt ttctgtgact ggtgagtagt caaccaagtc	2220
attctgagaa tagtgttatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacggataaa	2280
taccgcgcca catagcagaa cttaaaaagt gctcatcatt ggaaaacggtt cticggggcg	2340
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc	2400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtagcaaa aaacaggaag	2460
gcaaaatgcc gcaaaaaagg gaataaggc gacacggaaa tggtaatac tcatacttct	2520
ccttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt	2580
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc	2640
acctgacgcg ccctgttagcg gcgcattaag cgccgggggt gtggtggtta cgcgacgcgt	2700
gaccgctaca cttgccagcg ccctagcgcc cgctcccttc gcttcttcc ttcccttct	2760
cgccacgttc gccggcttcc cccgtcaagc tctaaatcgg gggctccctt tagggttccg	2820
atttagtgtt ttacggcacc tcgacccaa aaaacttgat tagggtgatg gttcacgtag	2880
tggccatcg ccctgtataga cggttttcg cccttgacg ttggagtcca cggttttaa	2940
tagtgactc ttgttccaaa ctggaaacaac actcaaccct atctcggtct attctttga	3000
tttataaggg attttgccga ttccggctt tggtaaaaa aatgagctga tttaacaaaa	3060
attnaacgcg aattnaaaca aaatattaac gtttacaatt tcccatcg cattcaggct	3120
gcgcaactgt tggttggc gatcggtgcg ggcctcttcg ctattacgcc agcccaagct	3180
accatgataa gtaagtaata ttaaggtacg ggaggtactt ggagcggccg caataaaaata	3240
tctttatattt cattacatct gtgtgttggt tttttgtgtg aatcgatagt actaacatac	3300
gctctccatc aaaacaaaac gaaacaaaac aaactagcaa aataggctgt ccccaagtcgca	3360
agtgcaggtg ccagaacatt tctctatcga taggtaccga gctcttacgc gtgcttagccc	3420

025CIP4 PCT SEQ List.txt

tcgagcagga tctatacatt	aatcaaatat tggcaatttag	ccatattagt cattggttat	3480
atagcataaa tcaatattgg	ctattggcca ttgcatacgt	tgtatctata tcataatatg	3540
tacatttata ttggctcatg	tccaatatga ccgccatgtt	gacattgatt attgactagt	3600
tattaatagt aatcaattac	ggggtcatta gttcatagcc	catatatgga gttccgcgtt	3660
acataactta cggtaaatgg	cccgcctggc tgaccgccc	acgaccccg cccattgacg	3720
tcaataatga cgtatgttcc	catagtaacg ccaataggga	cttccattg acgtcaatgg	3780
gtggagttt tacggtaaac	tgcccacttg gcagttacatc	aagtgtatca tatgccaagt	3840
ccgcccccta ttgacgtcaa	tgacggtaaa tggccgcct	ggcattatgc ccagtacatg	3900
accttacggg actttcctac	ttggcagttac atctacgtat	tagtcatcgc tattaccatg	3960
gtgatgcgtt tttggcagta	catcaatggg cgtggatagc	ggtttgactc acggggattt	4020
ccaagtctcc accccattga	cgtcaatggg agttgtttt	ggcaccaaaa tcaacgggac	4080
tttccaaaat gtcgtAACAA	ctccgcggca ttgacgcaaa	tggcggttag gcgtgtacgg	4140
tgggagggtct atataagcag	agctcgTTTA gtgaaccgtc	agatgcctg gagacgccat	4200
ccacgctgtt ttgacctcca	tagaagacac cgggaccgat	ccagcctccc ctcgaagctc	4260
gactctaggg gctcgagatc	cccgggtacc ggTCGCCACC	atggtgagca agggcgagga	4320
gctgttcacc ggggtggtgc	ccatcctggt cgagctggac	ggcgcacgtaa acggccacaa	4380
gttcagcgtg tccggcgagg	gcgagggcga tgccacctac	ggcaagctga ccctgaagtt	4440
catctgcacc accggcaagc	tgcccgtgcc ctggccacc	ctcgtgacca ccctgaccta	4500
cgcgctgcag tgcttcagcc	gctaccccga ccacatgaag	cagcacgact tcttcaagtc	4560
cgcctatgccc gaaggctacg	tccaggagcg caccatttc	ttcaaggacg acggcaacta	4620
caagacccgc gccgaggtga	agttcgaggg cgacaccctg	gtgaaccgca tcgagctgaa	4680
gggcatcgac ttcaaggagg	acggcaacat cctggggcac	aagctggagt acaactacaa	4740
cagccacaac gtctatatca	tggccgacaa gcagaagaac	ggcatcaagg tgaacttcaa	4800
gatccgccac aacatcgagg	acggcagcgt gcagctgcc	gaccactacc agcagaacac	4860
ccccatcgcc gacggccccc	tgctgctgcc cgacaaccac	tacctgagca cccagtccgc	4920
cctgagcaaa gaccccaacg	agaagcgcga tcacatggtc	ctgctggagt tcgtgaccgc	4980
cgccgggatc actctcggca	tggacgagct gtacaagtaa	agcggccgct cgagcatgca	5040
t			5041

<210> 7  
<211> 18116  
<212> DNA  
<213> Artificial Sequence  
<220>

## 025CIP4 PCT SEQ List.txt

&lt;223&gt; Plasmid p12.0lys-LSPIPNMM-CMV-pur-attB

<400> 7  
gggctgcagg aattcgattt ccgccttctt tgatattcac tctgttgtat ttcatctttt 60  
cttggcgtatg aaaggatata acagtctgtt taacagtctg tgaggaaata cttggtatttt 120  
cttctgtatca gtgttttat aagtaatgtt gaatatttga taaggctgtg tgtccttgt 180  
cttgggagac aaagcccaca gcaggtggtg gttggggtgg tggcagctca gtgacaggag 240  
aggttttttt gcctgtttttt tttttttttt ttttttttaa gtaaggtgtt ctttttctt 300  
agtaaaattttt ctactggact gtatgtttt acaggtcaga aacatttctt caaaagaaga 360  
accttttggaa aactgtacag ccctttctt tcattccctt tttgctttct gtgccaatgc 420  
ctttggttctt gattgcatta tggaaaacgt tgatcggAAC ttgaggtttt tatttatagt 480  
gtggcttgaa agcttggata gctgttgtt cacgagatac cttattaagt ttaggccagc 540  
ttgatgtttt atttttccccc tttgaagtag tgagcgttct ctggtttttt tcctttgaaa 600  
ctggtgaggc tttagattttt ctaatggat ttttacccgt atgatctagt tgcataccca 660  
aatgcttgta aatgttttcc tagttaacat gttgataact tcggatttac atgttgata 720  
tacttgcattt ctgtgtttctt agtaaaaata tatggcattt atagaaatac gtaattcctg 780  
atttccctttt tttttatctc tatgctctgt gtgtacaggt caaacagact tcactcctat 840  
tttttattttt agaattttt atgcagtctg tcggtgggttc ttgtgttgta aggatacagc 900  
cttaaatttc ctagagcgat gctcagtaag gcgggttgc acatgggttc aaatgtaaaa 960  
cgggcacgtt tggctgctgc cttcccgaga tccaggacac taaactgctt ctgcactgag 1020  
gtataaatcg cttagatcc cagggaaagtg cagatccacg tgcatttct taaagaagaa 1080  
tgaataactttt ctaaaatattt ttggcatagg aagcaagctg catggatttgc tttggactt 1140  
aaatttattttt ggtaacggag tgcataagggtt ttaaacacag ttgcagcatg ctaacgagtc 1200  
acagcgttta tgcagaagtg atgcctggat gcctgttgca gctgtttacg gcactgcctt 1260  
gcagtgagca ttgcagatag ggggtgggtg ctttgcgtcg tggccacaca cgctgccaca 1320  
cagccacccctc ccggaaacaca tctcacctgc tgggtacttt tcaaaccatc ttagcagtag 1380  
tagatgagtt actatgaaac agagaagttc ctcagttggaa tattctcatg ggatgtctt 1440  
tttcccatgt tggcaaaagt atgataaagc atctctattt gtaaattatg cacttggtag 1500  
ttcctgaatc ctttctatag caccacttgc tgcagcaggt gtggctctg gtgtggcctg 1560  
tgtctgtgct tcaatctttt aaagcttctt tggaaataca ctgacttgc tgaagtctct 1620  
tgaagatagt aaacagtact tacctttgat cccaatgaaa tcgagcattt cagttgtaaa 1680  
agaattccgc ctattcatac catgtaatgtt aattttacac ccccaagtgc gacactttgg 1740  
aatatattca agtaatagac tttggcctca ccctcttgc tactgtattt tgtaatagaa 1800

025CIP4 PCT SEQ List.txt  
aatattttaa actgtgcata tgattattac attatgaaag agacattctg ctgatcttca 1860  
aatgtaaaga aatgaggagt gcgttgctt ttataaatac aagtgattgc aaattagtgc 1920  
aggtgtccctt aaaaaaaaaaaa aaaaaaaagta atataaaaag gaccaggtgt tttacaagtg 1980  
aaatacattc ctatggta aacagttaca ttttatgaa gattaccagc gctgctgact 2040  
ttctaaacat aaggctgtat tgtcttcctg taccattgca tttcctcatt cccaatttgc 2100  
acaaggatgt ctggtaaac tattcaagaa atggcttga aatacagcat gggagcttgt 2160  
ctgagttgga atgcagagtt gcactgcaaa atgtcaggaa atggatgtct ctcagaatgc 2220  
ccaactccaa aggattttat atgtgtatat agtaagcagt ttcctgattc cagcaggcca 2280  
aagagtctgc tgaatgttgc gttgccggag acctgtatTT ctcaacaagg taagatggta 2340  
tcctagcaac tgccggatTT aatacattt cagcagaagt acttagttaa tctctacctt 2400  
tagggatcgt ttcatcattt ttagatgtta tacttgaaat actgcataaac ttttagctt 2460  
catgggttcc ttttttcag cctttaggag actgttaagc aatttgcgtt ccaacttttgc 2520  
tgttggtctt aaactgcaat agtagttac cttgtattga agaaataaaag accatttttgc 2580  
tattaaaaaa tactttgtc tgtcttcatt ttgacttgctc tgatatcctt gcagtgccta 2640  
ttatgtcagt tctgtcagat attcagacat caaaactttaa cgtgagctca gtggagttac 2700  
agctgcggtt ttgatgctgt tattatttct gaaactagaa atgatgttgc cttcatctgc 2760  
tcatcaaaca cttcatgcag agtgaaggc tagtgagaaa tgcatacatt tattgatact 2820  
tttttaaagt caactttta tcagattttt ttttcatttgc gaaatatattt gttttctaga 2880  
ctgcataagct tctgaatctg aaatgcagtc tgattggcat gaagaagcac agcactttc 2940  
atcttactta aacttcattt tggaaatgaaag gaagttaagc aagggcacag gtccatgaaa 3000  
tagagacagt gcgcctcagga gaaagtgaac ctggatttct ttggcttagtgc ttctaaatct 3060  
gtagtgagga aagtaacacc cgattccttgc aaagggtctcc agcttaatgc ttccaaatttgc 3120  
gaaggtggca ggcaacttgg ccactggta tttactgcat tatgtctcag tttcgccagct 3180  
aacctggctt ctccactatt gagcatggac tatagcctgg cttcagaggc caggtgaagg 3240  
ttgggatggg tggaaaggagt gctggctgt ggctgggggg actgtggggc ctccaaatct 3300  
agcttgggtt gggcagcaca gggaaaagtgc tggtaacta ttttaagta ctgtgttgc 3360  
aacgtctcat ctgcaaatac gtaggggtgt tactctcgaa gattaacagt gtgggttcag 3420  
taatataatgg atgaattcac agtggaaagca ttcaagggtt gatcatctaa cgacaccaga 3480  
tcatcaagct atgattggaa gcggtatcag aagagcgagg aaggtaaagca gtcttcataat 3540  
gttttcctc cacgtaaagc agtctggaa agtagcaccc cttgagcaga gacaaggaaa 3600  
taattcagga gcatgtgcta ggagaactttt cttgctgaat tctacttgca agagcttgc 3660  
tgcctggctt ctggtgccctt ctgcagcacc tgcaaggccc agagcctgtg gtgagctgg 3720

## 025CIP4 PCT SEQ List.txt

ggaaaagatt ctgctcaagt ccaagcttca gcaggtcatt gtcttgctt cttccccag	3780
cactgtcag cagagtggaa ctgatgtcga agcctcctgt ccactacctg ttgctgcagg	3840
cagactgctc tcagaaaaag agagctaact ctatgccata gtctgaaggt aaaatgggtt	3900
ttaaaaaaaga aaacacaaaag gcaaaaccgg ctgccccatg agaagaaagc agtggtaaac	3960
atggtagaaa aggtgcagaa gcccccaggc agtgtgacag gcccttcctg ccacctagag	4020
gcgggaacaa gcttccctgc ctagggctct gcccgcgaag tgcgtttc tttgggtgggt	4080
tttggggc gtttggtttt gagatttaga cacaaggaa gcctgaaagg aggtgttggg	4140
cactatttg gtttgtaaag cctgtacttc aaatatataat tttgtgaggg agtgtagcga	4200
attggccaat taaaaataaaa gttgcaagag attgaaggct gagtagttga gagggtaaca	4260
cgtttaatga gatcttctga aactactgct tctaaacact tgtttgagtg gtgagacett	4320
ggataggtga gtgctttgt tacatgtctg atgcacttgc ttgtcctttt ccatccacat	4380
ccatgcattc cacatccacg catttgcac ttatccata tctgtcatat ctgacatacc	4440
tgtctttcg tcacttggtc agaagaaaca gatgtgataa tccccagccg ccccaagttt	4500
gagaagatgg cagttgcttc tttccctttt tcctgctaag taaggatttt ctctggctt	4560
tgacacctca cgaaatagtc ttccctgcctt acattctggg cattatttca aatatcttg	4620
gagtgcgctg ctctcaagtt tgtgtcttcc tactcttaga gtgaatgctc tttagtgaa	4680
agagaaggaa gagaagatgt tggccgcagt tctctgatga acacacctct gaataatggc	4740
caaagggtggg tgggtttctc tgaggaacgg gcagcgtttgcctctgaaag caaggagctc	4800
tgcggagttt cagttatccc gcaactgatg gtggaaactgg tgcttaaagc agattcccta	4860
gttccctgc tacttctttt ctttcttggc agtcagtttta tttctgacag acaaacagcc	4920
accccccactg caggcttaga aagtatgtgg ctctgcctgg gtgtgttaca gctctgcct	4980
ggtgaaaggaa gattaaaacg ggcaccattc atcccaaaca ggatcctcat tcatggatca	5040
agctgttaagg aacttgggtt ccaacacctaa aacattaatt ggagtacgaa tggtaattaaa	5100
actgcattct cgcattccta agtcatttag tctggactct gcagcatgta ggtcggcagc	5160
tcccactttc tcaaagacca ctgatggagg agtagtaaaa atggagaccc attcagaaca	5220
accaacggag tggccgaa gaaactgatg gaaataatgc atgaattgtg tggtgacat	5280
tttttttaaa tacataaaact acttcaaattt aggtcggaga aggtcagtgt tttattagca	5340
gccataaaac caggtgagcg agtaccattt ttctctacaa gaaaaacgat tctgagctct	5400
gcgttaagtat aagttctcca tagcggctga agctcccccc tggctgcctg ccatctcagc	5460
tggagtgccag tgccatttcc ttggggtttc tctcacagca gtaatggac aataacttcac	5520
aaaaattctt tctttcctg tcatgtggaa tccctactgt gccctccctgg tttacgtta	5580

025CIP4 PCT SEQ List.txt

ccccctgact gttccattca gcggtttggaa aagagaaaaaa gaatttggaa ataaaacatg	5640
tctacgttat cacctcctcc agcattttgg ttttaatta tgtcaataac tggcttagat	5700
ttggaaatga gaggggggttg ggtgtattac cgaggaacaa aggaaggctt atataaactc	5760
aagtctttta ttttagagaac tggcaagctg tcaaaaacaa aaaggcctta ccaccaaatt	5820
aagtgaatag ccgctatagc cagcagggcc agcacgaggg atggtgact gctggcacta	5880
tgccacggcc tgcttgtac tctgagagca actgctttgg aaatgacagc acttggtgca	5940
atttcctttg tttcagaatg cgtagagcgt gtgcttggcg acagttttc tagttaggcc	6000
acttcctttt tccttctc ctcattctcc taagcatgtc tccatgctgg taatcccagt	6060
caagtgaacg ttcaaacaat gaatccatca ctgtaggatt ctcgtggta tcaaatcttt	6120
gtgtgaggta tataaaatat ggaagcttat ttatTTTcg ttcttccata tcagtcttct	6180
ctatgacaat tcacatccac cacagcaaat taaaggtgaa ggaggctggt gggatgaaga	6240
gggtcttcta gctttacgtt cttcccttgca aggccacagg aaaatgtga gagctgtaga	6300
atacagcctg gggtaagaag ttcagtctcc tgctggaca gctaaccgca tcttataacc	6360
ccttctgaga ctcatcttag gaccaaatacg ggtctatctg gggTTTTgt tcctgctgtt	6420
cctcctggaa ggctatctca ctatttact gctcccacgg ttacaaacca aagatacagc	6480
ctgaatTTTT tctaggccac attacataaa tttgacctgg taccaatatt gttctctata	6540
tagttatTC cttccccact gtgttaacc ccttaaggca ttcaagacaa ctagaatcat	6600
agaatggTTT ggatttggaaag gggcTTaaa catcatccat ttccaaccct ctgcccattgg	6660
ctgcttgcca cccactggct caggctgccc agggccccat ccagcctggc cttgagcacc	6720
tccagggatg gggcacccac agcttctctg ggcagcctgt gccaacacct caccactctc	6780
tggtaaaga attctttttt aacatctaatt ctaaatctct tcttttttag tttaaagcca	6840
ttcctctttt tcccgttgct atctgtccaa gaaatgtgta ttggtctccc tcctgcttat	6900
aagcaggaag tacttggagg ctgcagttag gtctccccac agcctctct tctccaggct	6960
gaacaagccc agctccttca gcctgtcttc gtaggagatc atcttagtgg ccctcctctg	7020
gaccCATTCC aacagttcca cggctttctt gtggagcccc aggtctggat gcagtacttc	7080
agatggggcc ttacaaaggc agagcagatg gggacaatcg cttacccctc cctgctggct	7140
gccccctgttt tgatgcagcc cagggtactg ttggcTTTC aggctccag accccttgc	7200
gatttgcgtc aagctttca tccaccagaa cccacgcttc ctggtaata cttctgccc	7260
cacttctgtta agcttgcTTTc aggagacttc cattctttag gacagactgt gttacaccta	7320
cctgccttat tcttgcataat atacattca gttcatgttt cctgttaacag gacagaatata	7380
gtattcctct aacaaaaata catgcagaat tccttagtgcc atctcagtag ggttttcatg	7440
gcagtattag cacatagtca atttgctgca agtaccttcc aagctgcggc ctcccataaa	7500

## 025CIP4 PCT SEQ List.txt

tcctgtat	ttt gggatcagtt acctttggg gtaagcttt gtatctgcag agaccctggg	7560
ggttctgat	g tgcttcagct ctgctctgtt ctgactgcac catttctag atcacccagt	7620
tgttcctgt	a caacttcctt gtcctccatc cttcccagc ttgtatctt gacaaataca	7680
ggcctat	ttt ttttgttgc tcagcagcca ttttaattctt cagtgtcatc ttgttctgtt	7740
gatgccact	g gaacaggatt ttcagcagtc ttgcaaagaa catctagctg aaaactttct	7800
gccattcaat	attcttacca gttcttctt tttgaggtga gccataaaatt actagaactt	7860
cgtcactgac	a aagtttatgc attttattac ttctattatg tacttacttt gacataaacac	7920
agacacgcac	atatttgct gggatttcca cagtgtctt gtgtccttca catggtttta	7980
ctgtcatact	tccgttataa ctttggcaat ctgcccagct gcccatcaca agaaaagaga	8040
ttcctttt	ttt attacttctc ttcagccaat aaacaaaatg tgagaagccc aaacaagaac	8100
ttgtggggca	ggctgcccattc aaggagaga cagctgaagg gttgtgttagc tcaatagaat	8160
taagaaataa	taaaagctgtg tcagacagtt ttgcctgatt tatacaggca ccccccaagc	8220
cagagaggct	gtctgccaag gccacccttgc agtccttggt ttgtaaagata agtcataggt	8280
aactttctg	gtgaatttgc tggagaatca tggatggcagt tcttgctgtt tactatggta	8340
agatgctaaa	ataggagaca gcaaagtaac acttgctgtc gttaggtgctc tgctatccag	8400
acagcgatgg	cactcgcaca ccaagatgag ggatgctccc agctgacgga tgctggggca	8460
gtaacagtgg	gtccccatgct gcctgctcat tagcatcacc tcagccctca ccagcccatc	8520
agaaggatca	tcccaagctg aggaaagttt ctcatcttct tcacatcatc aaacctttgg	8580
cctgactgat	gcctcccgga tgcttaatg tggtcactga catctttatt tttctatgat	8640
ttcaagtcag	aaacctccgga tcaggagggaa acacatagtg ggaatgtacc ctcagctcca	8700
aggccagatc	ttccttcaat gatcatgcat gctacttagg aagggtgtgt tttgtgaatg	8760
tagaattgcc	tttggttat tttcttcctg ctgtcaggaa catttgaat accagagaaa	8820
aagaaaagt	g ctcttcttgg catggagga gttgtcacac ttgcaaaata aaggatgcag	8880
tcccaaatgt	tcataatctc agggtctgaa ggaggatcag aaactgtgttata tacaatttca	8940
ggcttctctg	aatgcagctt ttgaaagctg ttccctggccg aggcagttact agtcagaacc	9000
ctcggaaaca	g gaaacaaatg tcttcaaggt gcagcaggag gaaacacctt gcccattcatg	9060
aaagtgaata	accactgccc ctgaaggaat ccagctcctg tttgagcagg tgctgcacac	9120
tcccacactg	aaacaacagt tcattttat aggacttcca ggaaggatct tcttcttaag	9180
cttcttaatt	atggtatcatc tccagttggc agatgactat gactactgac aggagaatga	9240
ggaactagct	g ggaatattt ctgtttgacc accatggagt cacccatttc tttactggta	9300
tttggaaata	ataattctga attgcaaagc aggagttgc gaagatcttca atttcttcca	9360

025CIP4 PCT SEQ List.txt  
tgttggtgac agcacagttc tggctatgaa agtctgctta caaggaagag gataaaaaatc 9420  
atagggataa taaaatctaag tttgaagaca atgagggttt agctgcattt gacatgaaga 9480  
aattgagacc tctactggat agctatggta tttacgtgtc ttttgctta gttacttatt 9540  
gaccccagct gaggtcaagt atgaactcag gtctctcggt ctactggcat ggattgatta 9600  
catacaactg taatttttagc agtgatttag ggttatgag tactttgca gtaaatcata 9660  
gggtagtta tgtaatctc agggaaaaaa aaaaaaagcc aaccctgaca gacatcccag 9720  
ctcaggtgga aatcaaggat cacagtcag tgcggtccca gagaacacag ggactcttct 9780  
cttaggacct ttatgtacag ggcctaaga taactgatgt tagtcagaag actttcattt 9840  
ctggccacag tttagctgag gcaatcctgg aattttctct ccgctgcaca gttccagtca 9900  
tcccagttt tacagttctg gcacttttg ggtcaggccg tgatccaagg agcagaagtt 9960  
ccagctatgg tcagggagtg cctgaccgtc ccaactcact gcactcaaac aaaggcgaaa 10020  
ccacaagagt ggctttgtt gaaattgcag tgtggccag aggggctgca ccagtactgg 10080  
attgaccacg aggcaacatt aatcctcagc aagtgcattt tgccaggcatt aaattgaact 10140  
aactgatact acaatgcaat cagtatcaac aagtggttt gcttggaa tggagtctag 10200  
gggctctaca ggagtagcta ctctctaattt gagttgcatt ttgaagcagg acactgtgaa 10260  
aagctggcct cctaaagagg ctgctaaaca ttagggtcaa tttccagtg cactttctga 10320  
agtgtctgca gttccccatg caaagctgcc caaacatagc acttccaatt gaatacaatt 10380  
atatgcaggc gtactgcttc ttgccagcac tgtccctctc aaatgaactc aacaaacaat 10440  
ttcaaagtct agtagaaagt aacaagcttt gaatgtcatt aaaaagtata tctgctttca 10500  
gtagttcagc ttatTTATGC ccactagaaa catcttgtac aagctgaaca ctggggctcc 10560  
agatttagtgg taaaacctac tttatacaat catagaatca tagaatggcc tgggttggaa 10620  
gggacccaa ggatcatgaa gatccaacac ccccgccaca ggcagggcca ccaacctcca 10680  
gatctggtac tagaccaggc agcccaggc tccatccaac ctggccatga acacctccag 10740  
ggatggagca tccacaacct ctctggcag cctgtgccag cacctcacca ccctcttgt 10800  
gaagaacttt tccctgacat ccaatctaag cttccctcc ttgaggtag atccactccc 10860  
ccttgtgcta tcactgtcta ctcttgaaa aagttgattc tcctccttt tggaaaggtag 10920  
caatgaggc tccttgagc cttctctct tctgcaggat gaacaagccc agctccctca 10980  
gcctgtctt ataggagagg tgctccagcc ctctgatcat ctggggcc ctccctctgga 11040  
cccgctccaa gagctccaca tcttcctgt actggggcc ccaggcctga atgcagtact 11100  
ccagatgggg cctcaaaaga gcagagaaa gagggacaat cacccctc accctgctgg 11160  
ccagccctct tctgatggag ccctggatac aactggcttt ctgagctgca acttctcctt 11220  
atcagttcca ctataaaac aggaacaata caacaggtgc tggatggccag tgcagagttt 11280

025CIP4 PCT SEQ List.txt

ttcacacttc ttcatttcgg tagatcttag atgaggaacg ttgaagttgt gcttctgcgt 11340  
gtgcttcttc ctcccaaata actcctgcct gatacctcac cccacctgcc actgaatggc 11400  
tccatggccc cctgcagcca gggccctgat gaaccggca ctgcttcaga tgctgtttaa 11460  
tagcacagta tgaccaagtt gcacccatga atacacaaac aatgtgttgc atccttcagc 11520  
acttgagaag aagagccaaa tttgcattgt cagggaaatgg tttagtaatt ctgccaatta 11580  
aaacttgttt atctaccatg gctgtttta tggctgttag tagtggtaca ctgatgatga 11640  
acaatggcta tgcagtaaaa tcaagactgt agatattgca acagactata aaattccct 11700  
gtggcttagc caatgtggta cttcccacat tgtataagaa atttggcaag tttagagcaa 11760  
tgtttgaagt gttggaaat ttctgtatac tcaagaggc gttttgaca actgtagaac 11820  
agaggaatca aaagggggtg ggaggaagtt aaaagaagag gcaggtgcaa gagagcttgc 11880  
agtccccctg tgtgtacgac actggcaaca tgaggtcttt gctaattttg gtgctttgct 11940  
tcctgcccct ggctgcctta gggcgatc tgcctcagac ccacagcctg ggcagcagga 12000  
ggaccctgat gctgctggct cagatgagga gaatcagcct gtttagctgc ctgaaggata 12060  
ggcacgattt tggcttcctt caagaggagt ttggcaacca gtttcagaag gctgagacca 12120  
tccctgtgct gcacgagatg atccagcaga tcttaacct gtttagcacc aaggatagca 12180  
gcfgctgcttggat ggtgagacc ctgctggata agtttacac cgagctgtac cagcagctga 12240  
acgatctggaa ggcttgcgtg atccaggcgc tggcgatc cgagaccctt ctgatgaagg 12300  
aggatagcat cctggctgtg aggaagtact ttcagaggat caccctgtac ctgaaggaga 12360  
agaagtacag cccctgcgt tggaaagtcg tgagggctga gatcatgagg agctttagcc 12420  
tgagcaccaa cctgcaagag agcttgaggt ctaaggagta aaaagtctag agtcggggcg 12480  
gccggccgt tcgagcagac atgataagat acattgatga gttggacaa accacaacta 12540  
gaatgcagtg aaaaaaatgc tttatttgcgaa ttatgtga tgctattgct ttatttgcgaa 12600  
ccattataag ctgcaataaa caagtttaca acaacaattt cattcatttt atgtttcagg 12660  
ttcaggggaa ggtgtggag gttttttaaa gcaagtaaaa cctctacaaa tgtggtaaaa 12720  
tcgataagga tccgtcgacc gatgcccttgc agagccttca acccagtcg ctcctccgg 12780  
tggcgccgg gcatgactat cgtcgccgca cttatgactg tcttctttat catgcaactc 12840  
gtaggacagg tgccggcagc gctcttcgc ttcctcgctc actgactcgc tgcgctcggt 12900  
cgttcggctg cggcgagcgg tatcagctca ctcaaaaggcg gtaatacggt tatccacaga 12960  
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 13020  
taaaaaggcc gcgttgcgtgg cgttttcca taggctccgc ccccctgacg agcatcacaa 13080  
aaatcgacgc tcaagtca ggtggcgaaa cccgacagga ctataaaagat accaggcg 13140

025CIP4 PCT SEQ List.txt  
tccccctgga agctccctcg tgcgtctcc tttccgacc ctggcgctta ccggataacct 13200  
gtccgcctt ctcccttcgg gaagcgtggc gcttctcaa tgctcacgct gtaggtatct 13260  
cagttcggtg taggtcggttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 13320  
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt 13380  
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc 13440  
tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggtat 13500  
ctgcgtctg ctgaagccag ttacccctgg aaaaagagtt ggtagctctt gatccggcaa 13560  
acaaccacc gctggtagcg gtggttttt tggcaag cagcagatta cgccgagaaa 13620  
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggAACGA 13680  
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 13740  
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggctctga 13800  
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc 13860  
catagttgcc tgactccccg tcgtgttagat aactacgata cgggagggct taccatctgg 13920  
ccccagtgct gcaatgatac cgccgagaccc acgctcaccg gctccagatt tatcagcaat 13980  
aaaccagcca gccggaaggg ccgagcgcag aagtggcct gcaactttat ccgcctccat 14040  
ccagtctatt aattgttgcc gggagctag agtaagttagt tcgcccgttta atagttgcg 14100  
caacgttgtt gccattgcta caggcatcg ggtgtcacgc tcgtcggtt gtagggcttc 14160  
attcagctcc gttcccaac gatcaaggcg agttacatga tccccatgt tgtgaaaaaa 14220  
agcggtagc tccttcggc ctccgatcg tgtcagaagt aagttggccg cagtgttatac 14280  
actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt 14340  
ttctgtgact ggtgagtagt caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 14400  
ttgctcttgc ccggcgtcaa tacggataa taccgcgcca catagcagaa cttaaaagt 14460  
gctcatcatt ggaaaacgtt cttcccccg aaaactctca aggatcttac cgctgtttag 14520  
atccagttcg atgtaaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac 14580  
cagcgttct gggtagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataaggc 14640  
gacacggaaa tggtaatac tcatactctt ctttttcaa tattattgaa gcatttatca 14700  
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 14760  
ggttccgcgc acattttccc gaaaagtgcc acctgacgcg ccctgttagcg ggcattaaag 14820  
cgccgggggt gtgggtgtta cgccgacgcgt gaccgctaca cttgccagcg ccctagcgcc 14880  
cgctcccttc gcttcttcc cttcccttct cgccacgttc gccggcttcc cccgtcaagc 14940  
tctaaatcgg gggctccctt tagggttccg atttagtgc ttacggcacc tcgaccccaa 15000  
aaaacttgat tagggtgatg gttcacgtag tggccatcg ccctgataga cggttttcg 15060

## 025CIP4 PCT SEQ List.txt

ccctttgacg ttggagtcca cgttcttaa tagtgactc ttgttccaaa ctggaacaac 15120  
actcaaccct atctcggtct attctttga tttataaggg atttgccga tttcgcccta 15180  
ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaca aaatattaac 15240  
gtttacaatt tcccattcgc cattcaggct gcgcactgt tggaaaggc gatcggtgcg 15300  
ggccttcg ctattacgcc agcccaagct accatgataa gtaagtaata ttaaggtacg 15360  
ggaggtactt ggagcggccg ctctagaact agtggatccc ccggccgcaa taaaatatct 15420  
ttatttcat tacatctgtg tggtggttt ttgtgtgaat cgatagtact aacatacgct 15480  
ctccatcaaa acaaaacgaa acaaaacaaa ctagcaaaat aggctgtccc cagtgcaga 15540  
gcaggtgcca gaacatttct ctatcgatag gtaccgagct cttacgcgtg ctagccctcg 15600  
agcaggatct atacattgaa tcaatattgg caattagcca tattagtcat tggttatata 15660  
gcataaatca atattggcta ttggccattt catacggtt atctatatca taatatgtac 15720  
atttatattt gctcatgtcc aatatgaccg ccatgttgc attgatttatt gactagttat 15780  
taatagtaat caattacggg gtcattagtt catagcccat atatggagtt ccgcgttaca 15840  
taacttacgg taaatggccc gcctggctga ccgcacaacg acccccgccc attgacgtca 15900  
ataatgacgt atgttcccat agtaacgcca atagggactt tccattgacg tcaatgggtg 15960  
gagtatttac ggtaaaactgc ccacttggca gtacatcaag tgtatcatat gccaaagtccg 16020  
ccccctattt acgtcaatga cggtaaatgg cccgcctggc attatgccc gtacatgacc 16080  
ttacggact ttccctacttg gcagtacatc tacgtatttag tcacgcstat taccatggtg 16140  
atgcggttt ggcagttacat caatggcggt ggatagcggt ttgactcacg gggatttcca 16200  
agtctccacc ccattgacgt caatggaggt ttgtttggc accaaaatca acgggacttt 16260  
ccaaaatgtc gtaacaactc cgccccattt acgcaaatgg gcggtaggcg tgtacgggtgg 16320  
gaggtctata taagcagagc tcgtttagtg aaccgtcaga tcgcctggag acgcaccca 16380  
cgctgtttt acctccatag aagacaccgg gaccgatcca gcctcccctc gaagctcgac 16440  
tctaggggct cgagatctgc gatctaagta agcttgcacg cctgcaggcggccacg 16500  
accggtgccg ccaccatccc ctgacccacg cccctgaccc ctcacaagga gacgaccttc 16560  
catgaccgag tacaagccca cggtgccct cgccacccgc gacgacgtcc cccggccgt 16620  
acgcacccctc gccgcccgt tcgcccacta ccccgccacg cgccacacccg tcgacccgga 16680  
ccgcccacatc gagcgggtca cggagctgca agaacttttc ctcacgcgcg tcgggctcga 16740  
catcggaag gtgtgggtcg cggacgacgg cgccgcggcgcg gccgtctggaa ccacgcccga 16800  
gagcgtcgaa gcggggccgg tgttcgccga gatcggcccg cgcatggccg agttgagcgg 16860  
ttcccggtcg gccgcccgtc aacagatgga aggccctcctg gcgcgcacc ggcccaagga 16920

025CIP4 PCT SEQ List.txt  
 gcccgcgtgg ttccctggcca ccgtcggcgt ctcgccccac caccaggggca agggtctggg 16980  
 cagcgccgtc gtgctccccg gagtgaggc ggccgagcgc gccgggtgc ccgccttcct 17040  
 ggagacctcc gcgccccgca acctcccctt ctacgagcgg ctcggcttca ccgtcaccgc 17100  
 cgacgtcgag gtgcccgaag gaccgcgcac ctggtgcatg acccgcaagc ccggtgccctg 17160  
 acgcccgcac cacgacccgc agcgccgcac cgaaaggagc gcacgacccc atggctccga 17220  
 ccgaagccga cccgggcggc cccgcccggc cccgcacccgc ccccgaggcc caccgactct 17280  
 agagtcgggg cggccggccg cttcgagcag acatgataag atacattgtat gagttggac 17340  
 aaaccacaac tagaatgcag tgaaaaaaat gctttatttg tgaattttgt gatgctattg 17400  
 ctttatttgt aaccattata agctgcaata aacaagtta caacaacaat tgcattcatt 17460  
 ttatgtttca gtttcagggg gaggtgtggg aggttttta aagcaagtaa aacctctaca 17520  
 aatgtggtaa aatcgataag gatcaattcg gtttcaggta ccgtcgcacga tgttaggtcac 17580  
 ggtctcgaag ccgcgggtcg ggtgccaggc cgtgcccctg ggctccccgg gcgcgtactc 17640  
 cacctcaccc atctggtcca tcatgatgaa cgggtcgagg tggcggtagt tgatcccgcc 17700  
 gaacgcgcgg cgacccggga agccctcgcc ctgcggaaaccg ctgggcgcgg tggtcacgg 17760  
 gagcacggga cgtgcgacgg cgtcgccggg tgccgatacg cggggcagcg tcagcgggtt 17820  
 ctcgacggtc acggcgggca tgcgacagc cgaattgatc cgtcgaccga tgcccttgag 17880  
 agccttcaac ccagttagt cttccgggtg ggccggggc atgactatcg tcgcccact 17940  
 tatgactgtc ttctttatca tgcaactcgt aggacaggtg ccggcagcgc tcttccgctt 18000  
 cctcgctcac tgactcgctg cgctcggtcg ttccggctgca gcgagcggta tcagctcact 18060  
 caaaggcggt aatacgggta tccacagaat cagggataa cgcaggaaag aacatg 18116

<210> 8  
 <211> 17402  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Plasmid pOMIFN-Ins-CMV-pur-attB

<400> 8  
 ggccgcacc accgtggagc tccaaattcgc cctatagtga gtcgtattac aattcactgg 60  
 ccgtcgaaaa acaacgtcgt gactggaaa accctggcgt tacccaaactt aatcgccctg 120  
 cagcacatcc cccttcgccc agctggcgtatagcgaaga ggccgcacc gatgcctt 180  
 cccaaacagtt ggcgcagcctg aatggcgaat gggacgcgcctt ctgttagcggc gcattaagcg 240  
 cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcctt ctagcgc 300  
 ctccttcgc tttcttcctt tccttcgc ccacgttcgc cggcttccc cgtcaagctc 360  
 taaatcgaaaa gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa 420

## 025CIP4 PCT SEQ List.txt

aacttgatta	gggtgatgg	tcacgttagt	ggccatcgcc	ctgatagacg	gttttcgcc	480
cttgacgtt	ggagtccacg	ttctttaata	gtggactctt	gttccaaact	ggaacaacac	540
tcaaccctat	ctcggtctat	tctttgatt	tataaggat	tttgcgatt	tcggcctatt	600
ggttaaaaaa	tgagctgatt	taacaaaaat	ttaacgcgaa	tttaacaaa	atattaacgc	660
ttacaattta	ggtggcactt	ttcggggaaa	tgtgcgcgga	accctattt	gtttatTTT	720
ctaaatACAT	tcaaatatgt	atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	780
atattgaaaa	aggaagagta	tgagtattca	acatttccgt	gtcgccctta	ttccctttt	840
tgccggcattt	tgccttcctg	ttttgctca	cccagaaacg	ctggtaaaag	taaaagatgc	900
tgaagatcag	ttgggtgcac	gagtgggtta	catcgaactg	gatctcaaca	gcggtaagat	960
ccttgagagt	tttcgccccg	aagaacgttt	tccaatgatg	agcacttttta	aagttctgct	1020
atgtggcgcg	gtattatccc	gtattgacgc	cgggcaagag	caactcggc	gccgcataca	1080
ctattctcag	aatgacttgg	ttgagtactc	accagtcaca	gaaaagcattc	ttacggatgg	1140
catgacagta	agagaattat	gcagtgctgc	cataaccatg	agtataaca	ctgcggccaa	1200
cttacttctg	acaacgatcg	gaggaccgaa	ggagctaacc	gctttttgc	acaacatggg	1260
ggatcatgta	actcgccctt	atcggtgg	accggagctg	aatgaagcca	taccaaacga	1320
cgagcgtgac	accacgatgc	ctgtagcaat	ggcaacaacg	ttgcgcaaac	tattaactgg	1380
cgaactactt	actctagctt	cccgcaaca	attaatagac	tggatggagg	cgataaaagt	1440
tgcaggacca	cttctgcgct	cggcccttcc	ggctggctgg	tttattgctg	ataaatctgg	1500
agccggtgag	cgtgggtctc	gcggtatcat	tgcaGactg	ggccagatg	gtaaGCCCTC	1560
ccgtatcgta	gttatctaca	cgacggggag	tcaggcaact	atggatgaac	gaaatagaca	1620
gatcgctgag	ataggtgcct	caTGattaa	gcattggtaa	ctgtcagacc	aagttaactc	1680
atataactt	tagattgatt	taaaacttca	tttttaattt	aaaaggatct	aggtgaagat	1740
cctttttgat	aatctcatga	ccaaaatccc	ttaacgtgag	tttcgttcc	actgagcgtc	1800
agaccccgta	gaaaagatca	aaggatctt	ttgagatcct	tttttctgc	gcgtaatctg	1860
ctgcttgcaa	acaaaaaaaaac	caccgctacc	agcggtggtt	tgttgccgg	atcaagagct	1920
accaactctt	tttccgaagg	taactggctt	cagcagagcg	cagataccaa	atactgtcct	1980
tctagtgtag	ccgtagttag	gccaccactt	caagaactct	gtacGaccgc	ctacatacac	2040
cgctctgcta	atccctgttac	cagtggctgc	tgccagtggc	gataagtcgt	gtcttaccgg	2100
gttggactca	agacgatagt	taccggataa	ggcgcagcgg	tcgggctgaa	cgggggggttc	2160
gtgcacacag	cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	2220
gctatgagaa	agcgccacgc	ttcccgaagg	gagaaaggcg	gacaggtatc	cggtaaGcg	2280

025CIP4 PCT SEQ List.txt  
cagggtcgga acaggagagc gcacgagggg gcttccaggg gaaacgcct ggtatcttta 2340  
tagtcctgtc gggtttcgccc acctctgact tgagcgtcga tttttgttat gctcgtcagg 2400  
ggggcggagc ctatggaaaa acgccagcaa cgccgcctt ttacggttcc tggcctttg 2460  
ctggccttt gctcacatgt tcttcctgc gttatcccct gattctgtgg ataaccgtat 2520  
taccgcctt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcggagtc 2580  
agtgagcggag gaagcggaaag agcgcccaat acgcaaaccg cctctccccc cgcgttggcc 2640  
gattcattaa tgcatgtggc acgacagggtt tcccgaactgg aaagcgggca gtgagcgc当地 2700  
cgcaattaat gtgagtttagc tcactcatta ggcaccccaag gctttacact ttatgcttcc 2760  
ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga 2820  
ccatgattac gccaagctcg aaattaaccc tcactaaagg gaacaaaagc tgggtaccgg 2880  
ccccccctc gactagaggg acagcccccc cccaaagccc ccagggatgt aattacgtcc 2940  
ctcccccgct agggggcagc agcgagccgc ccggggctcc gctccggtcc ggccgtcccc 3000  
ccgcatcccc gagccggcag cgtgcgggaa cagccgggc acggggagg tggcacggg 3060  
tcgcttcct ctgaacgctt ctcgctgctc tttgagcctg cagacacctg gggggatacg 3120  
ggaaaaaagc tttaggctga aagagagatt tagaatgaca gaatcataga acggcctggg 3180  
ttgcaaagga gcacagtgc catccagatc caacccctg ctatgtgcag ggtcatcaac 3240  
cagcagccca ggctgcccag agccacatcc agcctggcct tgaatgcctg cagggatggg 3300  
gcatccacag cctccttggg caacctgttc agtgcgtcac caccctctgg gggaaaaact 3360  
gcctcctcat atccaaccca aaccccccgtt ctctcagggtt aaagccattc ccccttgtcc 3420  
tatcaagggg gagttgctg tgacattgtt ggtctgggt gacacatgtt tgccaattca 3480  
gtgcatcactc gagaggcaga tcttgggat aaggaagtgc aggacagcat ggacgtggg 3540  
catgcaggtg ttgagggctc tgggacactc tccaagtcac agcggtcaga acagccttaa 3600  
ggataagaag ataggataga aggacaaaaga gcaagttaaa acccagcatg gagaggagca 3660  
caaaaaggcc acagacactg ctggccctg tgtctgagcc tgcattttt atgggtgtctg 3720  
gatgcaagca gaaggggtgg aagagcttgc ctggagagat acagctgggt cagtaggact 3780  
gggacaggca gctggagaat tgccatgttag atgttcatac aatcgtcaaa tcatgaaggc 3840  
tggaaaagcc ctccaagatc cccaaagacca accccaaacc acccaccgtg cccactggcc 3900  
atgtccctca gtgccacatc cccacagtcc ttcatcacct ccagggacgg tgacccccc 3960  
acctccgtgg gcagctgtgc cactgcagca ccgcctttt gagaaggtaa atcttgcataa 4020  
atccagcccg accctccctt ggcacaacgt aaggccattt tctctcatcc aactccagga 4080  
cgagtcagt gaggatgggg ctctagtcga ggtcgacggt atcgataagc ttgattaggg 4140  
agagcaatag gactctcaac ctcgtgagta tggcagcatg ttaactctgc actggagtcc 4200

## 025CIP4 PCT SEQ List.txt

agcgtggaa acaatctgcc ttgcacatga gtctcggtt gccaatattc cccaacggtt	4260
ttccttcagc ttgtctgtc tcctaagctc tcaaaacacc ttttggta ataaaactcac	4320
ttggcaacgt ttatctgtct taccttagtg tcacgttca tccctattcc ctttctcct	4380
cctccgtgtg gtacacagtg gtgcacactg gttctctgt tgatgttctg ctctgacagc	4440
caatgtgggt aaagttttc ctgccccgtg tctgtgttgt tttcacttca aaaaggccc	4500
tgggctcccc ttggagctct caggcatttc cttaatcatc acagtcacgc tggcaggatt	4560
agtccttcct aaaccttaga atgacctgaa cgtgtgctcc ctctttagt tcagtgcagg	4620
gagacgtttg cctcaagatc agggtccatc tcacccacag ggccattccc aagatgaggt	4680
ggatggtttta ctctcacaaa aagtttctt atgtttgct agaaaggaga actcactgcc	4740
tacctgtgaa ttcccctagt cctgggtctg ctgccactgc tgccctgtca gcctgtccc	4800
tggagggggc agcaactgct gtcacaaagg tgatcccacc ctgtctccac tgaaatgacc	4860
tcagtgccac gtgttgtata gggtataaag tacgggaggg ggatgcccgg ctcccttcag	4920
ggttgcagag cagaagtgtc tgtgtataga gtgtgtctt atctattaat gtaacagaac	4980
aacttcagtc ctatgtttt gtgggctgga attgcccattg tggtagggac aggccctgcta	5040
aatcactgca atcgcctatg ttctgaaggt atttggaaa gaaagggatt tgggggattt	5100
cctgtgattt gctttattt aatggcaaatt cacaggaaag cagttctgct caacagttgg	5160
ttgtttcagc caattcttgc agccaaagag ccgggtgccc agcgatataa tagttgtcac	5220
ttgtgtctgt atggatgaca gggaggtagg gtgacctgag gaccaccctc cagcttctgc	5280
tagcgttaggt acagtcacca cctccagctc cacacgagtc ccatcgttgtt ttaccaaaga	5340
aacacaatta tttggaccag tttggaaagt caccgcgtga attgtgaggc tagattaata	5400
gagctgaaga gcaaattttc ccaacttggaa gatactagtt ggtatttagta tcagaggaac	5460
agggccatag cacctccatg ctattagatt ccggctggca tgtacttttca aagatgattt	5520
gtaaactaaca atggcttatt gtgcttgc taagtctgtc tcctaattgtt aatgttcattt	5580
tggtttatatt aaccttcttgc ccatttgc ttcagggtttt cttgcagaac actggctgct	5640
ttaatcttagt ttaactgtttt ctgattattt cttaggata agatctgaat aaacttttttgc	5700
tggctttggc agacttttagc ttgggcttag ctcccacatt agcttttgc gcctttctg	5760
tgaagctatc aagatcctac tcaatgacat tagctgggtg caggtgtacc aaatcctgct	5820
ctgtggaaaca cattgtctga tgataccgaa ggcaaacgtg aactcaaaga ggcacagagt	5880
taagaagaag tctgtgcaat tcagaggaaa agccaaagtg gccatttagac acactttcca	5940
tgcagcattt gccagtaggt ttcatataaa actacaaaat ggaataaaacc actacaaaatg	6000
ggaaaagcct gatactagaa tttaaatatt cacccaggct caaggggtgt ttcatggagt	6060

025CIP4 PCT SEQ List.txt  
aatatcactc tataaaagta gggcagccaa ttattcacag acaaagctt ttttttctg 6120  
tgctgcagtg ctgttttcg gctgatccag ggtaacttat tgtgggtctg agagctgaat 6180  
gatttctcct tgtgtcatgt tggtgaagga gatatggca gggggagatg agcatgttca 6240  
agagggaaacg ttgcattttg gtggcttggg agaaaggtag aacgatatca ggtccatagt 6300  
gtcactaaga gatctgaagg atggttttac agaacagttg acttggctgg gtgcaggctt 6360  
ggctgtaaat ggatggaagg atggacagat gggtgacag agatttctgt gcaggagatc 6420  
atctcctgag ctcggtgctt gacagactgc agatccatcc cataaccttc tccagcatga 6480  
gagcgcgggg agctttggta ctgttcagtc tgctgcttgt tgcttcctgg gtgcacagtg 6540  
gtgattttct tactcacaca gggcaaaaac ctgagcagct tcaaagtgaa caggttgctc 6600  
tcataggcca ttcagttgtc aagatgaggt ttttggttc ttgttttcta aggtggaaag 6660  
aagcactgaa ggatcagttg cgagggcagg ggttagcac tgttcagaga agtcttattt 6720  
taactcctct catgaacaaa aagagatgca ggtgcagatt ctggcaagca tgcagtgaag 6780  
gagaaagccc tgaatttctg atatatgtgc aatgttggc acctaacatt ccccgctgaa 6840  
gcacagcagc tccagctcca tgcagttactc acagctggc cagccctcgg ctccagggtc 6900  
tgagcagtgc tgggactcac gaggttccat gtcttcaca ctgataatgg tccaatttct 6960  
ggaatgggtg cccatcctt gaggccccca aggccaggct ggctgcgtct ccgagcagcc 7020  
cgatctggcgtg gtgagtagcc agcccatggc aggagttaga gcctgatggt cttaaggtc 7080  
ccttccaacc taagccatcc tacgattcta ggaatcatga ctgtgagtg tgtattgcag 7140  
aggcaatatt ttaaagttat aaatgtttc tccccttcct tggttgc当地 agttatctt 7200  
atcgcccttat caatgctttt ggagtctcca gtcatttttca ttacamcaaa aagaggagga 7260  
agaatgaaga gaatcatttta atttcttgat tgaatagtag gattcagaaa gctgtacgta 7320  
atgccgtctc tttgtatcga gctgtaaggt ttctcatcat ttatcagcgt ggtacatatc 7380  
agcacttttc catctgatgt ggaaaaaaaaa atccttatca tctacagtct ctgtacctaa 7440  
acatcgctca gactcttac caaaaaagct ataggttttta aaactacatc tgctgataat 7500  
ttgcctgtt ttagctttc ttccatatgc tgcgtttgtg agaggtgcgt ggatggcct 7560  
aaactctcag ctgctgagct tgatgggtgc ttaagaatga agcactcact gctgaaaactg 7620  
tttcatttc acaggaatgt ttttagtggca ttgtttttat aactacatat tcctcagata 7680  
aatgaaatcc agaaataatt atgcaaactc actgcattccg ttgcacaggt ctatctgc 7740  
tagcaaagga aataatttg ggatggcaaa aacattcctt cagacatcta tatttaaagg 7800  
aatataatcc tggtacccac ccacttcatc cctcattatg ttcacactca gagataactca 7860  
ttctctgtt gttatcattt gatagcgtt tctttggttc tttgccacgc tctggctat 7920  
ggctgcacgc tctgcactga tcagcaagta gatgcgaggg aagcagcgt gagagggcgt 7980

## 025CIP4 PCT SEQ List.txt

gcccctcagct ggcacccagc cgctcagcct aggaggggac cttgccttc caccagctga	8040
ggtgcagccc tacaagctta cacgtgctgc gagcaggta gcaaaggag tcttcatggt	8100
gtgtttcttg ctgccccgaa gcaaaaacttt actttcattc attcccttig aagaatgagg	8160
aatgtttgga aacggactgc ttacgttca atttctctct tcccttaag gctcagccag	8220
gggccattgc tgaggacggc atcggggccc cctggaccaa atctgtggca cagatggtt	8280
cacttacatc agtggatgtg ggatctgcgc ctgtaatgtg tccttctgaa ggaaggaacg	8340
tgccttccaa gtgccagccc cacagccccc agccctccc tgtgctgctc caattcatct	8400
cctcttcctc cttctccctt tgctgttgt gctcggttag aaatcatgaa gattnagaag	8460
agaaaacaaa ataactggag tggaaaccca ggtgatgcag ttcattcagc tgtcataggt	8520
ttgtcggtgc tataggtctg tatcagagat gctarcacca ctttgctgac ggtgcttaac	8580
tcgggtgaac tctccttcac tcgcatttcatt tgcgggcctt atttacatcc ccagcatcca	8640
tcaccctctg ggaaaatggg cgcaactggat ctctaattgaa agactttccc tcttcagag	8700
cctgtggat gtgcagtgc acggaaacgtg gaggggctga gcagcagcac tgccccagg	8760
gagcaggagc ggatgccatc ggtggcagca tcccaaattga tgtcagcggta tgctgagcag	8820
gcagcggacg aacggacaga agcgatgcgt acaccttctg ttgacatggat atttggcagc	8880
gatttaacac tcgcttccta gtcctgctat tctccacagg ctgcattcaa atgaacgaag	8940
ggaagggagg caaaaagatg caaaatccga gacaaggcagc agaaatattt cttcgctacg	9000
gaagcgtgcg caaacaacct tctccaacag caccagaaga gcacagcgtaa cctttttca	9060
agaccagaaa agggaaattca caaaggctct gtggataacca gcgcgttcag ctctcctgat	9120
agcagatttc ttgtcagggtt gcgaatgggg tatggtgcca ggaggtgcag ggaccatatg	9180
atcatataca gcacagcagt cattgtgcat gtattaatat atattgagta gcagtgttac	9240
tttgc当地 caatagttca gagatgagtc ctgctgcata cctctatctt aaaactaact	9300
tataaatagt aaaaccttct cagttcagcc acgtgctcct ctctgtcagc accaatggtg	9360
cttcgc当地 acccagctgc aaggaatcag cccgtgatct cattaacact cagctctgca	9420
ggataaaatta gattgttcca ctctcttttgg ttgttaattt cgacggaaaca attgttca	9480
gctgatggtc ctaattgtca gctacagaaa acgtctccat gcagttcctt ctgcgc当地	9540
aaactgtcca ggctatagca ccgtgatgca tgctacctct cactccatcc ttcttctctt	9600
tcccaccagg gagagctgtg tggtttcaact ctcagccact ctgaacaata ccaaactgct	9660
acgcactgcc tccctcggaa agagaatccc cttgttgctt ttttatttac aggatccttc	9720
ttaaaaagca gaccatcatt cactgcaaac ccagagcttc atgcctctcc ttccacaacc	9780
gaaaacagcc ggcttcattt gtcttttta aatgctgttt tccaggtgaa ttttggccag	9840

025CIP4 PCT SEQ List.txt  
cgtgtggct gagatccagg agcacgtgtc agtttctgc ttcattgct cctgttctgc 9900  
attgcctctt tctggggttt ccaagagggg gggagactt gcgcggggat gagataatgc 9960  
ccctttctt agggtggctg ctggcagca gagtggtctt gggtaactgt ggcaccaatg 10020  
ggaggcacca gtgggggtgt gtttgtgca ggggggaagc attcacagaa tggggctgat 10080  
cctgaagctt gcagtccaag gcttgcgtc tgtacccagt gaaatccttc ctctgttaca 10140  
taaagcccag ataggactca gaaatgttagt cattccagcc ccccttcc tcagatctgg 10200  
agcagcactt gtttgcagcc agtcctcccc aaaatgcaca gacctcgccg agtggaggg 10260  
gatgtaaaca gcbaagggtta attacctcct tgtaaaaac actttgttgt ccatacatgt 10320  
ttctgtcaat cttaaaaac agaaccgaga ggcagcgagc actgaagagc gtgttccat 10380  
gctgagttaa tgagacttgg cagctcgctg tgccagatg atccctgtgc ttcatggag 10440  
gctgtaacct gtctccccat cgccatcaca ccgcagtgtc gtcctggaca cctcaccctc 10500  
cataagctgt aggtgcagc tgcccaggaa tcaagagact tttcctaagg ctcttaggac 10560  
tcatcttgc cgctcagtag cgtgcagcaa ttactcatcc caactatact gaatgggtt 10620  
ctgccagctc tgctgtttt tcaataagca tttcttcatt ttgcctctaa gtttctctca 10680  
gcagcaccgc tctgggtgac ctgagtggcc acctgaaacc cgaggggcac agccaccacc 10740  
tccctgttgc tgctgctcca gggactcatg tgctgctgga tggggggaaag catgaagttc 10800  
ctcacccaga cacctgggtt gcaatggctg cagcgtgctc ttcttggat gcagattgtt 10860  
tccagccatt actttagaa atgtgctgtg gaagccctt gtatctttt ctgtggccct 10920  
tcagcaaaag ctgtggaaa gctctgaggc tgcttctt ggtcgtggag gaattgtatg 10980  
ttccttctt aacaaaaatt atccttagga gagagcactg tgcaagcatt gtgcacataa 11040  
aacaattcag gttgaaaggg ctctctggag gtttccagcc tgactactgc tcgaagcaag 11100  
gccaggttca aagatggctc aggtgctgt gtgccttcct gattatctgt gccaccaatg 11160  
gaggagattc acagccactc tgctccctgt gccactcatg gagaggaata ttcccttata 11220  
ttcagataga atgttatcct ttagctcagc cttccctata accccatgag ggagctgcag 11280  
atccccatac tctcccttc tctgggtga aggccgtgtc ccccgcccc cttccacc 11340  
ctgtgcccta agcagccgc tggcctctgc tggatgtgtg cctatatgtc aatgcctgtc 11400  
cttgcagtcc agcctggac atttaattca tcaccagggt aatgtggAAC tgtgtcatct 11460  
tccctgcag ggtacaaagt tctgcacggg gtccttcgg ttcaggaaaa ctttcactgg 11520  
tgctacctga atcaagctct atttaataag ttcataagca catggatgtg tttccttaga 11580  
gatacgtttt aatggatca gtgattttta tttgctttgt tgcttacttc aaacagtgcc 11640  
tttggcagg aggtgaggga cgggtctgcc gttggctctg cagtgatttc tccaggcgtg 11700  
tggctcaggt cagatagtgg tcactctgtg gccagaagaa ggacaaagat ggaaattgca 11760

025CIP4 PCT SEQ List.txt

## 025CIP4 PCT SEQ List.txt

ggagatgtgg atgcagagac gaatgaagga cctgtccctt actcccccta gcattctgtg 13680  
ctattnnaggg ttctaccaga gtccttaaga gttttttttt ttttttggtc caaaagtctg 13740  
tttgggttgggt tttgaccact gagagcatgt gacacttgc tcaagctatt aaccaagtgt 13800  
ccagccaaaa tcaattgcct gggagacgca gaccattacc tggaggtcag gacctcaata 13860  
aatattacca gcctcattgt gccgctgaca gattcagctg gctgctccgt gttccagtc 13920  
aacagttcg acgccacgtt tgtatataatt tgcaaggcgc ctcgggggga ccatctcagg 13980  
agcagagcac cgccagccgc ctgcagagcc gggcagtacc tcaccatggc tttgaccctt 14040  
gccttactgg tggcttcctt ggtgctgagc tgcaagagca gctgctctgt gggctgcgt 14100  
ctgcctcaga cccacagcct gggcagcagg aggaccctga tgctgctggc tcagatgagg 14160  
agaatcagcc tggtagctg cctgaaggat aggacgattt tggctttcc tcaagaggag 14220  
tttggcaacc agttttagaa ggctgagacc atccctgtgc tgacagagat gatccagcag 14280  
atcttaacc tggtagcac caaggatagc agcgctgctt gggatgagac cctgctggat 14340  
aagtttaca ccgagctgta ccagcagctg aacgatctgg aggcttgcgt gatccagggc 14400  
gtggcgtga ccgagacccc tctgatgaag gaggatagca tcctggctgt gaggaagtac 14460  
tttcagagga tcaccctgta cctgaaggag aagaagtaca gcccctgcgc ttggaaagtc 14520  
gtgaggcgtg agatcatgag gagcttagc ctgagcacca acctgcaaga gagcttgg 14580  
tctaaggagt aaaaagtcta gagtcggggc ggccggccgc ttcgagcaga catgataaga 14640  
tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaaaatg ctttatttgt 14700  
gaaatttgt atgctattgc ttatattgta accattataa gctgcaataa acaagttAAC 14760  
aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtgggaa gttttttaa 14820  
agcaagtaaa acctctacaa atgtggtaaa atcgataccg tcgacctcga ctagagccgc 14880  
cactaacata cgctctccat caaaacaaaa cggaaacaaaa caaactagca aaataggctg 14940  
tccccagtgc aagtgcaggt gccagaacat ttctctatcg ataggtaccg agctcttacg 15000  
cgtgctagcc ctcgagcagg atctatacat tgaatcaata ttggcaatta gccatattag 15060  
tcattggta tatagcataa atcaatattt gctattggcc attgcatacg ttgtatctat 15120  
atcataatat gtacatttat attggctcat gtccaatatg accgcccgt tgacattgt 15180  
tattgactag ttatataatg taatcaatta cggggtcatt agttcatagc ccatatatgg 15240  
agttccgcgt tacataactt acggtaaatg gcccgcctgg ctgaccgcgg aacgacccc 15300  
gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt 15360  
gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc 15420  
atatgccaag tccgccccctt attgacgtca atgacggtaa atggccgc tggcattatg 15480  
cccagtagat gacccatgg gactttccta cttggcagta catctacgta ttagtcatcg 15540

## 025CIP4 PCT SEQ List.txt

ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact 15600  
cacggggatt tccaagtctc caccccatgg acgtcaatgg gagtttgaaa tggcacccaa 15660  
atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta 15720  
ggcgtgtacg gtgggaggtc tatataagca gagctcgaaa agtgaaccgt cagatcgcc 15780  
ggagacgcca tccacgctgt tttgacaccttcc atagaagaca ccgggaccga tccagcctcc 15840  
cctcgaagct cgactctagg ggctcgagat ctgcgatcta agtaagcttgc catgcctgca 15900  
ggtcggccgc cacgaccggc gcccggccacca tcccctgacc cacgcccctg acccctcaca 15960  
aggagacgac cttccatgac cgagtacaag cccacggtgc gcctcgccac ccgcgacgac 16020  
gtccccccggg ccgtacgcac cctcggccgc gcgttcggccg actaccccgcc cacgcgcccac 16080  
accgtcgacc cggaccgcca catcgagcgg gtcaccggc tgcaagaact cttcctcactg 16140  
cgcgtcgccgc tcgacatcggtt caagggtgtgg gtcgcggacg acggcgccgc ggtggcggtc 16200  
tggaccacgc cggagagcgt cgaagcgggg gcgggtttcg ccgagatcgcc cccgcgcatg 16260  
gccgagttga gcgggttcccg gctggccgcg cagcaacaga tggaaaggcct cctggccgc 16320  
caccggccca aggagccgcg gtggttcccg gccaccgtcg gcgtctcgcc cgaccaccag 16380  
ggcaagggtc tgggcagcgc cgtcgtgctc cccggagttgg aggcggccga gcgcgcccgg 16440  
gtgcccgcct tcctggagac ctccgcgcgc cgcacccctcc cttctacga gcggctcgcc 16500  
ttcacccgtca ccggccacgt cgaggtgcggc gaaggaccgc gcacctggc catgaccgc 16560  
aagcccggtg cctgacgcgc gccccacgc cccgcacgcg ccaccgaaag gagcgcacga 16620  
ccccatggct ccgaccgaaag ccgaccgggg cggcccccgc gacccgcac ccgccccgaa 16680  
ggcccaccga ctctagagtc ggggcggccg gccgcttcga gcagacatga taagatacat 16740  
tgatgatgtt ggacaaacca caactagaat gcagtaaaaaa aaatgtttta tttgtgaaat 16800  
ttgtgatgtt attgttttat ttgttaaccat tataagctgc aataaaacaag ttaacaacaa 16860  
caattgcatt cattttatgt tttaggttca gggggaggttgg tggaggtttttttaaagcaa 16920  
gtaaaacctc tacaaatgtg gtaaaatcga taaggatcaa ttccggcttca ggtaccgtcg 16980  
acgatgttagg tcacggtctc gaagccgcgg tgccgggtgcc agggcgtgcc cttggcgtcc 17040  
ccgggcgcgt actccacctc acccatctgg tccatcatga tgaacgggtc gaggtggcgg 17100  
tagttgatcc cggcgaacgc gcggcgcacc gggaaaggccct cgcgcgtcc accgctggc 17160  
gcgggttgtca cggtgagcac gggacgtgcg acggcgtcgg cgggtgcggta tacgcggggc 17220  
agcgtcagcg ggttctcgac ggtcacggcg ggcacgtcga cagccgaatt gatccgtcga 17280  
ccgatgcctt tgagagcctt caacccagtc agtccttcc ggtggccgcg gggcatgact 17340  
atcgtcgccg cacttatgac tgtcttcttt atcatgcaac tcgttaggaca ggtgcccggca 17400

## 025CIP4 PCT SEQ List.txt

gc

17402

<210> 9  
<211> 5172  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Plasmid pRSV-Int

<400> 9  
ctgcattaat gaatcgccca acgcgcgggg agaggcggtt tgcgtattgg gcgctttcc 60  
gcttcctcgc tcactgactc gctgcgctcg gtcgttccgc tgcggcgagc ggtatcagct 120  
caactcaaagg cggttaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 180  
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgttttc 240  
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 300  
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgtct 360  
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaaagcgtg 420  
gcgcttctc aatgctcacg ctgttaggtat ctcagttcgg tgttaggtcgt tcgctccaag 480  
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatac cggttaactat 540  
cgtcttgagt ccaacccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac 600  
aggattagca gagcgaggtt tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 660  
tacggctaca ctagaaggac agtatttggt atctgcgttc tgctgaagcc agttaccttc 720  
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggttt 780  
tttggggca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 840  
ttttctacgg ggtctgacgc tcagtgaaac gaaaactcac gttaaggat tttgggtcatg 900  
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 960  
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 1020  
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgctgtgt 1080  
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 1140  
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccgaaag ggccgagcgc 1200  
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccggaaagct 1260  
agagtaagta gttcgccagt taatagtttgc caacacgttgc ttgccattgc tacaggcatc 1320  
gtgggtgtcac gctcgctgtt tggtatggct tcattcagct ccggttccca acgtcaagg 1380  
cgagttacat gatccccat gttgtcaaa aaagcggtta gctccttcgg tcctccgatc 1440  
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 1500  
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 1560

## 025CIP4 PCT SEQ List.txt

tcattctgag aatagtgtat gcggcgaccg agttgcttt gcccggcgtc aatacgggat	1620
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg	1680
cggaaactct caaggatctt accgctgtt agatccagtt cgatgttaacc cactcgtgca	1740
cccaactgat cttagcattc ttttactttc accagcgtt ctgggtgagc aaaaacagga	1800
aggcaaaatg ccgaaaaaaa ggaaataagg ggcacacgga aatgttgaat actcatactc	1860
ttccttttc aatattattt aagcattt cagggattt gtctcatgag cgatatacata	1920
tttgaatgta tttagaaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg	1980
ccacctgacg tcgacggatc gggagatctc ccgatcccc atggtcgact ctcagtacaa	2040
tctgctctga tgccgcatac ttaagccagt atctgctccc tgcttggtg ttggaggtcg	2100
ctgagtagtg cgcgagcaaa attaaagcta caacaaggca aggcttgacc gacaattgca	2160
tgaagaatct gcttagggtt aggctttt cgctgcttcg cgatgtacgg gccagatata	2220
cgctgtcttagg gggcttagga tcgattctag gaattctcta gccgcgtct agggatcccg	2280
gcgcgtatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatct	2340
gctccctgct tgtgtgttgg aggtcgctga gtatgcgcg agcaaaaattt aagctacaac	2400
aaggcaaggc ttgaccgaca attgcatgaa gaatctgctt agggtaggc gtttgcgc	2460
gcttcgcgat gtacgggcca gatatacgat tatctgaggg gactagggtg tgtttagg	2520
aaaagcgggg ctccgggtt acgcggtag gagtccccctc agatatagt agttcgctt	2580
ttgcataggg agggggaaat gtagtcttat gcaatacact tgttagtctt caacatggta	2640
acgatgagtt agcaacatgc cttacaagga gaaaaaaagc accgtgcatt ccgattgg	2700
gaagtaaggt ggtacgatcg tgccttatta ggaaggcaac agacaggtct gacatggatt	2760
ggacgaacca ctgaattccg cattgcagag ataattgtat ttaagtgcct agtcgatac	2820
aataaacgcc atttgaccat tcaccacatt ggtgtgcacc tccaagctt catgcctgca	2880
ggtaccggtc cggattccc gggtcgcacg gctcaactgt cgtagggtcg ccgacatgac	2940
acaagggtt gtgaccgggg tggacacgta cgcgggtgct tacgaccgatc agtcgcgcga	3000
gcgcgagaat tcgagcgcag caagcccagc gacacagcgt agcgcacacg aagacaaggc	3060
ggccgacctt cagcgcgaag tcgagcgcga cggggccgg ttcaggttcg tcgggcattt	3120
cagcgaacgc cggggcacgt cggcggttcgg gacggcggag cgccggagt tcgaacgcatt	3180
cctgaacgaa tgccgcgcgc ggcggctcaa catgatcatt gtctatgacg tgtcgccctt	3240
ctcgccctg aaggtcatgg acgcgattcc gattgtctcg gaattgctcg ccctggcgt	3300
gacgattgtt tccactcagg aaggcgtctt cccgcaggaa aacgtcatgg acctgattca	3360
cctgattatg cggctcgacg cgtcgacaa agaatcttcg ctgaagtcgg cgaagattct	3420

## 025CIP4 PCT SEQ List.txt

cgacacgaag aaccttcagc gcgaattggg cgggtacgtc ggccggaaagg cgcccttacgg	3480
cttcgagctt gtttcggaga cgaaggagat cacgcgcaac ggccgaatgg tcaatgtcgt	3540
catcaacaag ctgcgcact cgaccactcc ccttaccgga cccttcgagt tcgagcccga	3600
cgtaatccgg tggtggtggc gtgagatcaa gacgcacaaa caccttccct tcaagccggg	3660
cagtcaagcc gccattcacc cgggcagcat cacggggctt tgtaagcgca tggacgctga	3720
cgccgtgccg acccggggcg agacgattgg gaagaagacc gcttcaagcg cctgggaccc	3780
ggcaaccgtt atgcgaatcc ttcgggaccc gcgtattgcg ggcttcgccc ctgaggtgat	3840
ctacaagaag aagccggacg gcacgcccac cacgaagatt gagggttacc gcattcagcg	3900
cgacccgatc acgctccggc cggtcgagct tgattgcgga ccgatcatcg agcccgctga	3960
gtggtatgag cttcaggcgt gtttgacgg cagggggcg gcgaaggggc ttccccggg	4020
gcaagccatt ctgtccgcca tggacaagct gtactgcgag tgtggcgccg tcatgacttc	4080
gaagcgcggg gaagaatcga tcaaggactc ttaccgctgc cgtcgccgga aggtggtcga	4140
cccggtccgca cctgggcagc acgaaggcac gtcaacgac agcatggcgg cactcgacaa	4200
gttcgttgcg gaacgcacatct tcaacaagat caggcacgac gaaggcgacg aagagacgtt	4260
ggcgcttctg tggaaagccg cccgacgctt cggcaagctc actgaggcgc ctgagaagag	4320
cggcgaacgg gcgaaccttg ttgcggagcg cggcagcgc ctgaacgccc ttgaagagct	4380
gtacgaagac cgcgcggcag gcgcgtacga cggaccggtt ggcaggaagc acttccggaa	4440
gcaacaggca gcgcgtacgc tccggcagca aggggcggaa gagcggcttg ccgaacttga	4500
agccgcccggaa gccccgaagc ttcccccttga ccaatggttc cccgaagacg ccgacgctga	4560
cccgaccggc cctaagtcgt ggtggggcg cgcgtcagta gacgacaagc gcgtgttcgt	4620
cgggctcttc gtagacaaga tcgttgtcac gaagtcgact acgggcaggg ggcagggaac	4680
gcccattcgag aagcgcgctt cgatcacgtg ggcgaagccg ccgaccgacg acgacgaaga	4740
cgacgcccag gacggcacgg aagacgtacgc ggcgtacgca gacacccggta tccctcgagg	4800
ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct	4860
tctagttgcc accatctgt tttttgtcccc tccccgtgc cttccttgcac cctggaaaggt	4920
gccactccca ctgtcctttc ctaataaaat gaggaaatttgcatcgatttgc tctgagtagg	4980
tgtcattcta ttctgggggg tgggtgggg caggacagca agggggagga ttggaaagac	5040
aatagcaggc atgctggggta tgccgtgggc tctatggctt ctgaggcgga aagaaccagg	5100
tgcccagtca tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtcaatcc	5160
actggggcg cg	5172

<210> 10  
<211> 6233

## 025CIP4 PCT SEQ List.txt

<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Plasmid pCR-XL-TOPO-CMV-pur-attB

&lt;400&gt; 10

agcgcccaat acgcaaaccg cctctccccg cgcggtggcc gattcattaa tgcagctggc 60  
acgacagggtt tcccgaactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagtttagc 120  
tcactcatta ggcacccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa 180  
tttgtagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat 240  
tttaggtgacg cgtagaata ctcaagctat gcatcaagct tggtaaccgag ctcggatcca 300  
ctagtaacgg ccgccagtgt gctggaattc gcccttggcc gcaataaaaat atctttattt 360  
tcattacatc tgggtgtgg tttttgtgt gaatcgatag tactaacata cgctctccat 420  
caaaaacaaaaa cgaaacaaaaa caaactagca aaataggctg tccccagtgc aagtgcaggt 480  
gccagaacat ttctctatcg ataggtaaccg agctcttacg cgtgctagcc ctcgagcagg 540  
atctatacat tgaatcaata ttggcaatta gccatattag tcattggta tatagcataa 600  
atcaatattt gctattggcc attgcatacg ttgtatctat atcataatat gtacatttat 660  
attggctcat gtccaatatg accgcccattgt tgacatttgat tattgactag ttattaatag 720  
taatcaatta cggggtcatt agttcatagc ccatatatgg agttccgcgt tacataactt 780  
acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac gtcaataatg 840  
acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg ggtggaggtat 900  
ttacggtaaa ctgcccactt ggcagtagat caagtgtatc atatgccaag tccgccttct 960  
attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtagat gaccttacgg 1020  
gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat ggtgatgcgg 1080  
ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt tccaagtctc 1140  
cacccattt acgtcaatgg gagtttggtt tggcacaaaa atcaacggga ctttccaaaa 1200  
tgtcgtaaca actccgccttcc attgacgcaa atgggcggta ggcgtgtacg gtggggaggtc 1260  
tatataagca gagctcggtt agtgaaccgt cagatcgctt ggagacgcca tccacgctgt 1320  
tttgacctcc atagaagaca ccgggaccga tccagcctcc cctcgaagct cgactctagg 1380  
ggctcgagat ctgcgatcta agtaagcttg catgcctgca ggtcgccgc cacgaccgg 1440  
gccgccacca tcccctgacc cacgcccctg accccctcaca aggagacgac cttccatgac 1500  
cgagtacaag cccacgggtgc gcctcgccac cccgcacgac gtcccccggg ccgtacgcac 1560  
cctcgccgc gcgttcgccc actaccccgac cacgcgcccac accgtcgacc cggaccgcca 1620  
catcgagcgg gtcaccgagc tgcaagaact cttcctcactg cgcgtcgccg tcgacatcgg 1680

025CIP4 PCT SEQ List.txt  
caaggtgtgg gtcgcggacg acggcgcgc ggtggcggtc tggaccacgc cggagagcgt 1740  
cgaagcgaaa gcggtgttcg ccgagatcgg cccgcgcatt gccagttga gcggttccc 1800  
gctggccgcg cagcaacaga tggaaggcct cctggcgcgg caccggccca aggagcccgc 1860  
gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag ggcaagggtc tggcagcgc 1920  
cgtcgtgctc cccggagtgg aggccggccga gcgcgcggg gtgcccgcct tcctggagac 1980  
ctccgcgcgc cgcacactcc cttctacga gcggctcgcc ttcacccgtca ccgcgcacgt 2040  
cgaggtgccc gaaggaccgc gcacctggtg catgaccgc aagccgggtg cctgacgccc 2100  
gccccacgac cgcacgcgc cgaccgaaag gagcgcacga ccccatggct ccgaccgaag 2160  
ccgaccgggg cgccccgcgc gacccgcac ccgcgcgcga ggcccaccga ctctagagtc 2220  
ggggcgccg gccgcttcga gcagacatga taagatacat ttagtggatggatggacaaacca 2280  
caactagaat gcagtgaaaa aaatgtttt tttgtgaaat ttgtgtatgtt attgttttat 2340  
ttgttaaccat tataagctgc aataacaacaa ttaacaacaa caattgcatt cattttatgt 2400  
ttcaggttca gggggagggtg tgggagggtt tttaaagcaa gtaaaacctc tacaaatgtg 2460  
gtaaaatcga taaggatcaa ttcggcttca ggtaccgtcg acgatgttagg tcacggtctc 2520  
gaagccgcgg tgccgggtgcc agggcgtgcc cttgggctcc ccggcgcgt actccacctc 2580  
acccatctgg tccatcatga tgaacgggtc gaggtggcgg tagttgtatcc cggcgaacgc 2640  
gcggcgcacc gggaaaggccct cgcgcgcggaa accgctggc gcggcggtca cggtagcac 2700  
gggacgtgcg acggcgtcgg cgggtgcgg tacgcggggc agcgtcagcg ggttctcgac 2760  
ggtcacggcg ggcgttcga cagccgaaatt gatccgtcga ccgtatgcct tgagagcctt 2820  
caacccagtc agctccttcc ggtggcgcg gggcatgact atcgtcgcgc cacttatgac 2880  
tgtcttcttt atcatgcaac tcgttaggaca ggtgccggca gcgccttcg cttccctcg 2940  
tcactgactc gctgcgttcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 3000  
cggttaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg aagggcgaat 3060  
tctgcagata tccatcacac tggccgcgc tcgagcatgc atctagaggg cccaattcgc 3120  
cctatagtga gtcgtattac aattcactgg ccgtcgaaaa acaacgtcgt gactggaaa 3180  
accctggcgt tacccaaattt aatgccttgc cagcacatcc cccttcgcgc agctggcgta 3240  
atagcgaaga ggccgcacc gatgcgcctt cccaaacagtt gcgcgccta tacgtacggc 3300  
agtttaaggt ttacacctat aaaagagaga gccgttatcg tctgtttgtg gatgtacaga 3360  
gtgatattat tgacacgcgg gggcgcacggc tggtgtatccc cctggccagt gcacgtctgc 3420  
tgtcagataa agtctccgt gaaaccttacc cggtggtgca tatcggggat gaaagctggc 3480  
gcatgatgac caccgatatg gccagtgtgc cggtctccgt tatcggggaa gaagtggctg 3540  
atctcagcca ccgcgaaaat gacatcaaaa acgcattaa cctgatgttc tggggaatat 3600

## 025CIP4 PCT SEQ List.txt

aaatgtcagg catgagatta tcaaaaagga tcttcaccta gatcctttc acgtagaaaag	3660
ccagtccgca gaaacggtgc tgaccccgga tgaatgtcag ctactggct atctggacaa	3720
gggaaaacgc aagcgaaag agaaagcagg tagcttcag tggcttaca tggcgatagc	3780
tagactggc ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgcctctg	3840
gtaagggtgg gaagccctgc aaagtaaact ggatggctt ctcggccca aggatctgat	3900
ggcgcagggg atcaagctct gatcaagaga caggatgagg atcgttcgc atgattgaac	3960
aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc ggctatgact	4020
gggcacaaca gacaatcggc tgctctgatg ccgcgtgtt ccgcgtgtca gcgcaggggc	4080
ccccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg caagacgagg	4140
cagcgcggct atcgtggctg gccacgcgg gcgttccctg cgacgtgtc ctcgacgttg	4200
tcaactgaagc gggaaaggac tggctgtat tggcgaagt gccggggcag gatctcctgt	4260
catctcacct tgctcctgcc gagaaagtat ccatcatggc tcatgcattg cggcggctgc	4320
atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcag	4380
cacgtactcg gatggaaagcc ggtctgtcg atcaggatga tctggacgaa gagcatcagg	4440
ggctcgcgcc agccgaactg ttgcgcaggc tcaaggcag catgcccac ggcgaggatc	4500
tcgtcgtgac ccatggcgat gcctgctgc cgaatatcat ggtggaaaat ggccgccttt	4560
ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg	4620
ctacccgtga tattgctgaa gagcttggcg gcgaatggc tgaccgcctc ctcgtcctt	4680
acggtatcgc cgctcccgat tcgcacgc tgccttcta tcgccttctt gacgagttct	4740
tctgaattat taacgcttac aatttcctga tgcggatttt tctccttacg catctgtcg	4800
gtatttcaca ccgcatacag gtggcactt tcggggaaat gtgcgcggaa cccctatttg	4860
tttatttttc taaatacatt caaatatgtt tccgctcatg agacaataac cctgataaaat	4920
gcttcataaa tagcacgtga ggaggccac catggccaag ttgaccagt ccgttccgg	4980
gctcaccgcg cgacgtcg ccggagcggt cgagttctgg accgaccggc tcgggttctc	5040
ccgggacttc gtggaggacg acttcgcgg tgtggccgg gacgacgtga ccctgttcat	5100
cagcgcggtc caggaccagg tggtgccga caacaccctg gcctgggtgt gggtgccgg	5160
cctggacgag ctgtacgcgg agtggtcgga ggtcgtgtcc acgaacttcc gggacgcctc	5220
cgggccggcc atgaccgaga tcggcgagca gccgtgggg cggagttcg ccctgcgcga	5280
cccgccggc aactgcgtgc acttcgtggc cgaggagcag gactgacacg tgctaaaact	5340
tcattttaa tttaaaagga tcttaggtgaa gatcctttt gataatctca tgacaaaaat	5400
cccttaacgt gagtttcgt tccactgagc gtcagacccc gtagaaaaaga tcaaaggatc	5460

025CIP4 PCT SEQ List.txt

ttcttgagat ccttttttc tgcgctaat ctgctgcttgc caaacaaaaaa aaccaccgct	5520
accagcggtg gtttgtttgc cgatcaaga gctaccaact cttttccga aggtaactgg	5580
cttcagcaga gcgcagatac caaatactgt cttcttagtg tagccgtagt taggccacca	5640
cttcaagaac tctgttagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc	5700
tgctgccagt ggcgataagt cgtgtttac cgggttggac tcaagacgat agttaccgga	5760
taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac	5820
gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga	5880
agggagaaaag gccggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag	5940
ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg	6000
acttgagcgt cgattttgt gatgctcgta agggggcgg agcctatgga aaaacgccag	6060
caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca tggctttcc	6120
tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttagtctgataccgc	6180
tcgcccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aag	6233

<210> 11  
<211> 234  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> attP containing polynucleotide

<400> 11  
gacttagtact gacggacaca ccgaagcccc ggcggcaacc ctcagcgat gccccggggc  
ttcacgtttt cccaggtcag aagcggtttt cgggagtagt gcccccaactg gggtaacctt  
ttagttctct cagttggggg cgtagggtcg ccgacatgac acaaggggtt gtgaccgggg  
tggacacgta cgcgggtgct tacgaccgac agtcgcgcga ggcgcactag taca  
234

<210> 12  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer attB-for

<400> 12  
taccgtcgac gatgttaggtc acggtc  
26

<210> 13  
<211> 11  
<212> PRT  
<213> SV40

<400> 13

025CIP4 PCT SEQ List.txt  
Cys Gly Gly Pro Lys Lys Lys Arg Lys Val Gly  
1 5 10

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US2005/006072

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, Sequence Search

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	THYAGARAJAN B ET AL: "Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase." MOLECULAR AND CELLULAR BIOLOGY. JUN 2001, vol. 21, no. 12, June 2001 (2001-06), pages 3926-3934, XP002330999 ISSN: 0270-7306 the whole document	1-142
Y	WO 03/024199 A (AVIGENICS, INC; RAPP, JEFFREY, C; SUTRAVE, PRAMOD) 27 March 2003 (2003-03-27) the whole document	1-142
Y	WO 03/025146 A (AVIGENICS, INC; RAPP, JEFFREY, C; CHRISTMANN, LEANDRO) 27 March 2003 (2003-03-27) the whole document	1-142
	-/-	

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

9 June 2005

Date of mailing of the international search report

06/07/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Novak-Giese, S

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US2005/006072

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HARVEY AJ ET AL.: "Expression of exogenous protein in the egg white of transgenic chicken" NATURE BIOTECHNOLOGY, vol. 19, April 2002 (2002-04), pages 396-399, XP002973268 the whole document -----	1-142
A	WO 99/19472 A (UNIVERSITY OF GEORGIA RESEARCH FOUNDATION) 22 April 1999 (1999-04-22) the whole document -----	1-142
A	IVARIE R: "Avian transgenesis: progress towards the promise" TRENDS IN BIOTECHNOLOGY, ELSEVIER PUBLICATIONS, CAMBRIDGE, GB, vol. 21, no. 1, January 2003 (2003-01), pages 14-19, XP002983884 ISSN: 0167-7799 the whole document -----	
P,X	WO 2004/092351 A (AVIGENICS, INC) 28 October 2004 (2004-10-28) the whole document -----	1-142
P,X	WO 2004/080162 A (AVIGENICS, INC) 23 September 2004 (2004-09-23) the whole document -----	1-142
P,X	WO 2004/067713 A (AVIGENICS INC; UNIVERSITY OF GEORGIA RESEARCH FOUNDATION INC; IVARIE,) 12 August 2004 (2004-08-12) the whole document -----	1-142

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2005/006072

### Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: 1-142 (partially)  
because they relate to subject matter not required to be searched by this Authority, namely:  
The subject-matter of claims 1-142 (partially) could not be searched since it relates also to transgenic humans and the modification of the germ line of humans. The claims should be restricted to 'non-human' vertebrates.
2.  Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

#### Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No

PCT/US2005/006072

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
WO 03024199	A	27-03-2003	WO	03024199 A2	27-03-2003
WO 03025146	A	27-03-2003	EP	1438401 A2	21-07-2004
			WO	03025146 A2	27-03-2003
			US	2003126629 A1	03-07-2003
			EP	1478751 A2	24-11-2004
			WO	02079447 A2	10-10-2002
			US	2003140363 A1	24-07-2003
WO 9919472	A	22-04-1999	AU	752946 B2	03-10-2002
			AU	1189999 A	03-05-1999
			CA	2307840 A1	22-04-1999
			EP	1023442 A1	02-08-2000
			JP	2001520009 T	30-10-2001
			WO	9919472 A1	22-04-1999
			US	6730822 B1	04-05-2004
			US	2004019922 A1	29-01-2004
			US	2004019923 A1	29-01-2004
			US	2004158882 A1	12-08-2004
WO 2004092351	A	28-10-2004	US	2004210954 A1	21-10-2004
			US	2004255345 A1	16-12-2004
			WO	2004092351 A2	28-10-2004
			US	2005034186 A1	10-02-2005
			US	2005066383 A1	24-03-2005
			WO	2004080162 A2	23-09-2004
			WO	2005040215 A2	06-05-2005
			US	2004226057 A1	11-11-2004
			US	2005003414 A1	06-01-2005
WO 2004080162	A	23-09-2004	US	2004210954 A1	21-10-2004
			US	2004255345 A1	16-12-2004
			US	2005034186 A1	10-02-2005
			US	2005066383 A1	24-03-2005
			WO	2004080162 A2	23-09-2004
			WO	2004092351 A2	28-10-2004
			WO	2005040215 A2	06-05-2005
WO 2004067713	A	12-08-2004	US	2004019922 A1	29-01-2004
			US	2004019923 A1	29-01-2004
			WO	2004067713 A2	12-08-2004